

**Modèles de prédiction de krachs boursiers
issus de bulles économiques**

par

Eric CHARLAND

mémoire présenté au Département de mathématiques
en vue de l'obtention du grade de maître ès sciences (M.Sc.)

FACULTÉ DES SCIENCES
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, février 2011



Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-79814-0
Our file *Notre référence*
ISBN: 978-0-494-79814-0

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Le 18 février 2011

*le jury a accepté le mémoire de Monsieur Éric Charland
dans sa version finale.*

Membres du jury

Professeur Thomas Brustle
Directeur de recherche
Département de mathématiques

Professeur Bernard Colin
Membre
Département de mathématiques

Professeur Éric Marchand
Président rapporteur
Département de mathématiques.

À l'humanité, qui aura toujours au moins ses propres limites à ignorer...

SOMMAIRE

L'effondrement des marchés boursiers est un phénomène de plus en plus fréquent et affecte plus ou moins directement la population entière de la planète. Il est donc primordial de développer des outils afin de modéliser et de prévoir de tels événements.

Mais pourquoi alors n'avons-nous pas encore proposé de solution ultime à un problème aussi important ? Quelles sont les problématiques principales qui nous en empêchent ? C'est dans cet ordre d'idées que nous exposerons divers modèles proposés par des experts de référence dans leurs domaines respectifs et que nous identifierons les diverses problématiques qui surviennent lors de l'étude de ces modèles.

Une étude plus approfondie de l'un des modèles actuels les plus présomptueux sera effectuée : il s'agit du modèle LPPL (ou *Log periodic power-law*), suggéré pour la première fois en 1996 par Sornette, Johansen et Bouchaud [60]. Ces derniers suggèrent notamment que les prix d'une action en présence d'une bulle économique soient modélisés selon la formule :

$$\ln p(t) = A + B(t_c - t)^\beta \{1 + C \cos(\omega \ln(t_c - t) + \phi)\}.$$

Enfin, un dernier chapitre sera consacré à la correction d'une prépublication récemment soumise à arXiv par Lin et Sornette [32] qui propose deux nouveaux modèles purement stochastiques. Note intéressante : le premier de ces modèles englobe sous forme de cas particulier la formule de Black-Scholes.

REMERCIEMENTS

Je tiens d'abord à remercier mon directeur de maîtrise, M. Thomas Brüstle, pour m'avoir proposé ce merveilleux domaine d'étude que sont les mathématiques financières ainsi que pour les nombreux échanges mathématiques qui m'ont été si bénéfiques dans cet apprentissage. Je tiens également à souligner la disponibilité de M. Mario Lambert que j'ai interrompu à maintes reprises dans des contextes très diversifiés ; ses précieux conseils de \LaTeX auront d'ailleurs été essentiels à l'élaboration de ce document.

Mes plus cordiaux remerciements sont adressés à M. David Brée qui, de l'autre côté de l'océan Atlantique et sans me connaître, aura accepté de répondre à toutes les questions que je lui ai posées concernant les travaux de Sornette *et al.* Sa contribution aura été cruciale à l'aboutissement de ce projet.

Merci à ma famille, à ma douce moitié, ainsi qu'à mes amies et amis, qui m'auront épaulé et encouragé tout au long de cette aventure. Je vous aime, discrètement mais sincèrement.

Ce mémoire a été rendu possible grâce au soutien financier du Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG).

Eric Charland
Sherbrooke, novembre 2010

TABLE DES MATIÈRES

SOMMAIRE	v
REMERCIEMENTS	vii
TABLE DES MATIÈRES	viii
LISTE DES TABLEAUX	xiii
LISTE DES FIGURES	xv
INTRODUCTION	1
CHAPITRE 1 — Matière préalable	3
1.1 Mouvement brownien	3
1.2 Martingale	10
1.3 Intégrale de Lebesgue	11
1.4 Calcul d'Itô	15
1.4.1 Intégrale d'Itô	15
1.4.2 Processus d'Itô	17

1.4.3	Formule d'Itô-Doeblin	18
1.5	Black-Scholes	21
1.5.1	Modèle de Black-Scholes	22
1.5.2	Formule de Black-Scholes	26
CHAPITRE 2 — Modèle classique		29
2.1	Un peu d'histoire	29
2.1.1	Brown et le mouvement brownien	29
2.1.2	Bachelier, Einstein et la marche aléatoire	30
2.1.3	Wiener et son processus	31
2.1.4	De Bachelier à Black-Scholes	31
2.2	Le modèle classique et la réalité	34
2.2.1	Données aberrantes et queues lourdes	34
2.2.2	L'indépendance des mouvements et l'amnésie des marchés	42
2.3	Les hypothèses du marché	45
2.3.1	Hypothèse d'efficience du marché	45
2.3.2	Anticipations rationnelles	46
2.3.3	Autres modèles	47
CHAPITRE 3 — Modèles contemporains		49
3.1	Problèmes inhérents à l'étude des marchés	49
3.1.1	Qu'est-ce que le risque ?	49
3.1.2	Comportement interne ou dynamique globale ?	51

3.1.3	Autorégulation du marché	52
3.1.4	L'unicité des données	53
3.1.5	Qu'est-ce qu'une tendance	54
3.1.6	Qu'est-ce qu'une bulle économique	56
3.2	À la recherche d'une distribution appropriée	56
3.2.1	Continuité du marché	57
3.2.2	Engle et ARCH	60
3.2.3	Mandelbrot et ses multiples modèles	63
3.2.4	David X. Li et les copules	67
CHAPITRE 4 — Modèle LPPL		71
4.1	D'où vient le modèle	72
4.1.1	Obtention de la formule	73
4.2	Tests empiriques	81
4.2.1	Les sept paramètres et leurs contraintes	81
4.2.2	Estimation des paramètres	82
4.2.3	Sornette <i>et al.</i>	86
4.2.4	Brée et Joseph	87
4.2.5	Expérimentations	88
4.3	Anomalies et problématiques	99
4.3.1	Hypothèses non vérifiables	100
4.3.2	Données brutes ou logarithme ?	100

4.3.3	Martingale et l'hypothèse des anticipations rationnelles	101
4.3.4	Sept paramètres et le surajustement	102
4.3.5	Critères d'identification d'un krach	102
4.3.6	Sensibilité de l'ajustement	104
4.3.7	Limitations pratiques	105
4.4	Efforts et développements récents	106
4.5	Discussion	107
CHAPITRE 5 — Modélisation stochastique		109
5.1	Premier modèle	110
5.1.1	Cas simplifié	110
5.1.2	Cas général	111
5.2	Second modèle	114
5.2.1	Cas simplifié	114
5.2.2	Cas général	116
CONCLUSION		119
BIBLIOGRAPHIE		121

LISTE DES TABLEAUX

4.1	Contraintes sur les paramètres LPPL	82
4.2	Sensibilité aux vecteurs initiaux — Dow Jones (1987)	91
4.3	Sensibilité à la fenêtre temporelle — Dow Jones (1987)	92
4.4	Ajustement au Dow Jones (1987) — Données brutes	94
4.5	Ajustement aux prix de l'or (USD) — Logarithme	96
4.6	Ajustement aux prix de l'or (USD) — Données brutes	98

LISTE DES FIGURES

1.1	Marche aléatoire symétrique	4
1.2	Marche aléatoire symétrique renormalisée (n=5)	5
1.3	Réalisation d'un mouvement brownien	9
1.4	Intégrale de Lebesgue — Partitionnement	12
1.5	Intégrale de Lebesgue — Discontinuités	13
2.1	Krach d'octobre 1987 (« Black Monday »)	36
2.2	Mouvements du Dow Jones (%) — Octobre 1928 à mai 2010	38
2.3	Dow Jones — Octobre 1928 à mai 2010 — 500 classes	39
2.4	Dow Jones — Octobre 1928 à mai 2010 — 1500 classes	39
2.5	Dow Jones — Octobre 1928 à mai 2010	40
2.6	Dow Jones — Septembre 2008 à mai 2010	40
2.7	Diagramme de Henry (Dow Jones) — Octobre 1928 à mai 2010	42
2.8	Variations (%) du Dow Jones — 1928 à 2010	43
2.9	Variations (%) du Dow Jones (zoom) — 1928 à 2010	44
2.10	Variations (% , valeur absolue) du Dow Jones — Janvier 1990 à mai 2010	44

3.1	FOREX — Une journée	58
3.2	FOREX — Janvier 2004 à mars 2008	58
3.3	Chute du Dow Jones — 6 mai 2010	59
3.4	Dow Jones — Logarithme des classes d’incréments positifs	64
4.1	Oscillations log-périodiques précédant un krach	72
4.2	Réseau hiérarchique en diamant (3 itérations)	77
4.3	Paramètres LPPL	80
4.4	Meilleurs ajustements au Dow Jones — Krach d’octobre 1987	94
4.5	Meilleurs ajustements aux prix de l’or	98

INTRODUCTION

Dans notre société, rien n'a davantage d'influence que l'argent : les entreprises cherchent constamment à maximiser leur profit, les décisions politiques sont prises en fonction de l'état actuel ou anticipé de l'économie et même les recherches universitaires sont orientées en fonction du financement disponible. Seules les entreprises, les institutions financières et une minime portion de la population s'intéressent aux variations quotidiennes de la bourse, mais en temps de crise, la planète entière en ressent les conséquences.

Avec l'effondrement récent des marchés boursiers à travers le monde et étant donnée la fréquence accrue de tels événements depuis un certain nombre d'années, il n'a jamais été aussi important de s'y attarder. Bien qu'il puisse sembler impensable d'empêcher de tels événements, il est toutefois possible de développer des outils mathématiques servant à suivre l'évolution d'un marché et, dans le meilleur des mondes, en arriver à prévoir l'arrivée d'un krach.

Les institutions financières utilisent à ce jour des outils développés il y a près de 40 ans et la majorité des institutions académiques enseignent encore cette théorie sans mentionner les problèmes rattachés à son application. Parmi ces outils se trouve la formule bien connue de Black-Scholes qui est utilisée encore aujourd'hui afin de déterminer le prix de diverses options sur les marchés boursiers. Celle-ci suppose entre autres des incréments — soit la différence entre deux prix consécutifs d'un même actif à la bourse — indépendants et normalement distribués, deux hypothèses qui sont réfutées par bon nombre de chercheurs depuis maintenant un quart de siècle. Alors que ce modèle s'avère raisonnable la majorité du temps,

il est évident qu'il est inadéquat lorsque le marché devient plus chaotique, en particulier précédent un krach qui est issu d'une bulle spéculative. Or, c'est spécifiquement durant ces périodes qu'un modèle fiable est indispensable.

Et alors si le modèle de Black-Scholes est inadéquat en périodes mouvementées, quelles alternatives s'offrent à nous ? Quelles sont les problématiques liées à l'élaboration d'un tel modèle ? De Louis Bachelier aux plus récents développements sur le sujet, nous survolerons dans ce document les efforts de plusieurs chercheurs de renommée au cours du dernier siècle.

Quoique les notions de base en calcul stochastique et en mathématiques financières nécessaires à la compréhension de ce mémoire soient reprises depuis le début, nous supposons du lecteur une connaissance approfondie en probabilité et en théorie de la mesure, notamment concernant les σ -algèbres, les processus stochastiques et les espérances conditionnelles.

CHAPITRE 1

Matière préalable

Dans ce chapitre, nous établirons les bases des sujets nécessaires à la compréhension des sections ultérieures. Plusieurs concepts et résultats sont empruntés de [53].

1.1 Mouvement brownien

Afin de définir le *mouvement brownien* (également appelé *processus de Wiener* en mathématiques), il est intéressant de considérer d'abord les marches aléatoires.

Définition 1.1.1

Soit ω_i une épreuve de Bernoulli avec issues équiprobables P (pile) et F (face) — c'est-à-dire le lancer d'une pièce de monnaie équilibrée. Posons $\omega = \omega_1\omega_2\omega_3 \dots$ la suite infinie de telles épreuves où ω_n correspond à l'issue de la $n^{\text{ème}}$ épreuve. Définissons

$$X_j = \begin{cases} 1 & \text{si } \omega_j = P, \\ -1 & \text{si } \omega_j = F. \end{cases} \quad (1.1)$$

Posons enfin $M_0 = 0$ et

$$M_k = \sum_{j=1}^k X_j, \quad k = 1, 2, \dots \quad (1.2)$$

Alors le processus $\{M_k\}_{k=0,1,2,\dots}$ est une **marche aléatoire symétrique** (unidimensionnelle).

La marche aléatoire $\{M_k\}_{k \in \mathbb{N}}$ (que nous noterons abusivement M_k) est donc un processus discret à intervalles réguliers qui augmente ou diminue d'une unité à chaque étape. Nous avons l'habitude de représenter graphiquement une réalisation d'une marche aléatoire en reliant M_k à M_{k+1} à l'aide d'un segment de droite pour $k = 0, 1, 2, \dots$

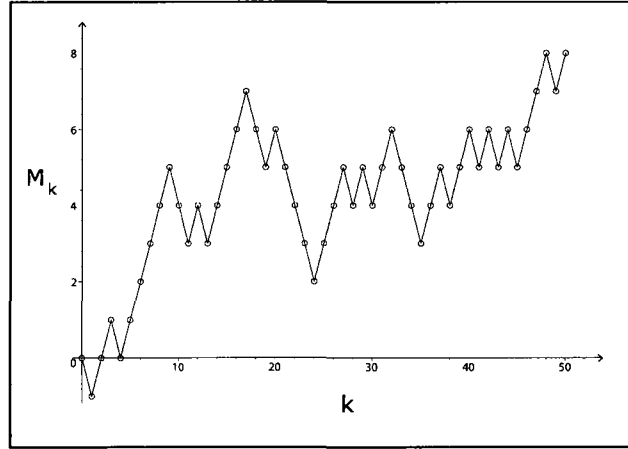


Figure 1.1 : Marche aléatoire symétrique

Définition 1.1.2

La **variation quadratique** d'une marche aléatoire symétrique M_k **jusqu'au temps** n , notée $[M, M]_n$, est la valeur définie par :

$$[M, M]_n = \sum_{j=1}^n (M_j - M_{j-1})^2 \quad (1.3)$$

Nous vérifions aisément que $[M, M]_n = n$.

Définition 1.1.3

Soit M_k une marche aléatoire symétrique telle que donnée par la définition 1.1.1. Fixons un entier positif n . Pour $t \geq 0$ tel que nt soit un entier, posons

$$W^{(n)}(t) = \frac{1}{\sqrt{n}} M_{nt}. \quad (1.4)$$

Alors $W^{(n)}(t)$ est appelé une **marche aléatoire symétrique renormalisée**.

Comme dans le cas d'une marche aléatoire symétrique, nous effectuerons une interpolation linéaire entre deux points consécutifs d'une réalisation d'une marche aléatoire symétrique renormalisée afin de la représenter graphiquement.

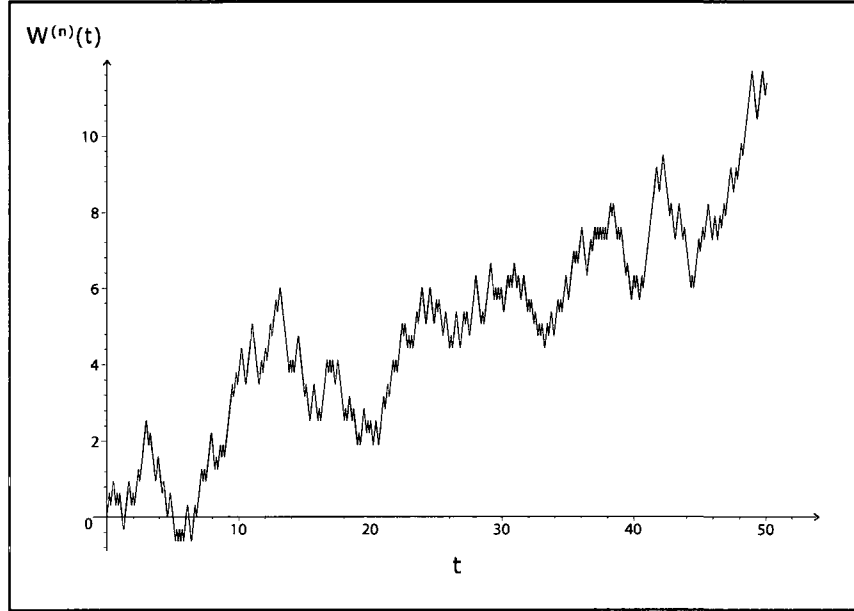


Figure 1.2 : Marche aléatoire symétrique renormalisée (n=5)

L'équation (1.4) nous dit qu'à tout temps $t > 0$ tel que nt soit entier, la marche aléatoire symétrique renormalisée n'est en fait qu'une marche aléatoire symétrique qui a subi une contraction de ses axes. En d'autres termes, les « augmentations et diminutions d'une unité » à chaque étape sont en fait des « augmentations ou diminutions de $\frac{1}{\sqrt{n}}$ unités » et lesdites « étapes » surviennent n fois plus souvent (ou rapidement) qu'au cours de la marche aléatoire symétrique.

Proposition 1.1.4

La variation quadratique d'une marche aléatoire symétrique renormalisée est donnée par :

$$[W^{(n)}, W^{(n)}](t) = \sum_{j=1}^{nt} \left[W^{(n)}\left(\frac{j}{n}\right) - W^{(n)}\left(\frac{j-1}{n}\right) \right]^2 = t \quad (1.5)$$

pour $t \geq 0$ tel que nt soit entier.

Proposition 1.1.5

Soit $W^{(n)}(t)$ une marche aléatoire symétrique renormalisée. Fixons $t \geq 0$. Alors la distribution de $W^{(n)}(t)$ évaluée au temps t converge en loi vers une distribution normale de moyenne zéro et de variance t lorsque $n \rightarrow \infty$.

Remarques 1.1.6 - Posons $X(t) := \lim_{n \rightarrow \infty} W^{(n)}(t)$. Ainsi, pour $0 < t_i < t_j < t_k < t_l$, nous avons :

1. $X(t_j) - X(t_i)$ suit une loi normale ;
2. $X(t_j) - X(t_i)$ et $X(t_l) - X(t_k)$ sont indépendants ;
3. $\mathbb{E}[X(t_j) - X(t_i)] = 0$;
4. $\text{Var}[X(t_j) - X(t_i)] = t_j - t_i$.

En constatant que $[X, X](t) = [\lim_{n \rightarrow \infty} W^{(n)}, \lim_{n \rightarrow \infty} W^{(n)}](t)$, nous déduisons le résultat suivant de la proposition 1.1.4. Les détails sont disponibles dans [53].

Corollaire 1.1.7

Soit $X(t) = \lim_{n \rightarrow \infty} W^{(n)}(t)$ où $W^{(n)}(t)$ est une marche aléatoire symétrique renormalisée. La variation quadratique de X au temps $t > 0$ est

$$[X, X](t) = t. \quad (1.6)$$

Afin de définir ce que nous entendons par « processus stochastique », nous aurons besoin des définitions suivantes empruntées de la théorie de la mesure.

Définition 1.1.8

Un ensemble Ω muni d'une σ -algèbre \mathcal{F} , noté (Ω, \mathcal{F}) , est appelé **ensemble mesurable**.

Définition 1.1.9

Un espace mesurable (Ω, \mathcal{F}) muni d'une mesure de probabilité \mathbb{P} , noté $(\Omega, \mathcal{F}, \mathbb{P})$, est appelé **espace probabilisé**.

Définition 1.1.10

Soient (Ω, \mathcal{F}) et (Ω', \mathcal{F}') deux espaces mesurables. Une fonction $f : \Omega \rightarrow \Omega'$ est dite $\mathcal{F} / \mathcal{F}'$ -**mesurable** si $f^{-1}(B) := \{\omega \in \Omega \mid f(\omega) \in B\} \subseteq \mathcal{F}$ pour tout $B \in \mathcal{F}'$. Lorsque $\Omega' \subseteq \mathbb{R}$, nous dirons simplement que f est \mathcal{F} -**mesurable** (ou encore **mesurable**).

Pour nos besoins, nous considérerons dorénavant uniquement les fonctions à valeurs réelles, c'est-à-dire $\Omega' \subseteq \mathbb{R}$. Dans le cas où Ω' est un intervalle de la droite réelle, nous prendrons pour acquis que \mathcal{F}' est la σ -algèbre de Borel correspondante.

Définition 1.1.11

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Une fonction \mathcal{F} -mesurable $X : \Omega \rightarrow \mathbb{R}$ est appelée **variable aléatoire**.

Définition 1.1.12

Soit (Ω, \mathcal{F}) un espace mesurable. Une suite $\{\mathcal{F}_t\}_{t \geq 0}$ de sous- σ -algèbres de \mathcal{F} est appelée **filtration de \mathcal{F}** si $\mathcal{F}_1 \subseteq \mathcal{F}_2$ pour $t_1 \leq t_2$.

Définition 1.1.13

Un **processus stochastique** est une collection de variables aléatoires $\{X_t\}_{t \in T}$ définies sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ où T est un ensemble d'indices. Si T est un intervalle, nous dirons qu'il s'agit d'un processus en temps continu ; sinon, nous dirons que le processus est discret. La notation $\{X(t)\}_{t \in T}$ est parfois utilisée plutôt que $\{X_t\}_{t \in T}$.

Définition 1.1.14

Soit $\{X_t\}_{t \in T}$ un processus stochastique et soit $F_X(x_{t_1}, \dots, x_{t_k})$ la fonction de répartition conjointe de $\{X_t\}$ aux temps t_1, \dots, t_k . Le processus $\{X_t\}$ est dit **stationnaire** si pour tous k, τ et t_1, \dots, t_k , nous avons :

$$F_X(x_{t_1+\tau}, \dots, x_{t_k+\tau}) = F_X(x_{t_1}, \dots, x_{t_k}).$$

Définition 1.1.15

Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $\{\mathcal{F}_t\}_{t \in T}$ une filtration de \mathcal{F} . Un processus stochastique $\{X_t\}_{t \in T}$ sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ est dit **adapté** à la filtration $\{\mathcal{F}_t\}_{t \in T}$ si X_t

est \mathcal{F}_t -mesurable pour tout $t \in T$.

Nous supposons dorénavant que tous les processus stochastiques que nous abordons sont adaptés à une filtration ; nous n'entrerons toutefois pas dans les détails de la filtration en question.

Nous pouvons enfin définir ce que nous entendons par « mouvement brownien ».

Définition 1.1.16

Un **mouvement brownien** (ou processus de Wiener) est un processus stochastique $\{W(t)\}_{t \geq 0}$ sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ tel que :

(i) $W(0) = 0$;

(ii) si $0 \leq t_0 \leq t_1 \leq \dots \leq t_k$ et H_i est un borélien pour $1 \leq i \leq k$, alors

$$\mathbb{P}[W(t_i) - W(t_{i-1}) \in H_i, 1 \leq i \leq k] = \prod_{1 \leq i \leq k} \mathbb{P}[W(t_i) - W(t_{i-1}) \in H_i];$$

(iii) si $0 \leq s < t$, alors l'incrément $W(t) - W(s)$ suit une loi normale de moyenne zéro et de variance $t - s$.

En d'autres termes, le mouvement brownien est tout simplement un processus stochastique dont les incréments sont indépendants, stationnaires et normalement distribués.

Remarque 1.1.17 - Dans la définition 1.1.16, la notation $W(t)$ a été utilisée plutôt que celle proposée précédemment, soit $W_t(\omega)$. Au besoin, la valeur du processus en ω sera notée $W(t, \omega)$. Ce choix de notation s'explique par l'utilisation que nous ferons du processus : alors que Ω correspond à l'ensemble des réalisations possibles du processus — et donc $\omega \in \Omega$ correspond à une réalisation spécifique —, $W(t, \cdot)$ est la trajectoire d'une réalisation du processus dans le temps. En considérant une unique réalisation du processus, il est donc inutile de spécifier la valeur de ω .

L'existence de processus stochastiques $\{W(t)\}_{t \geq 0}$ satisfaisant les critères de la définition

1.1.16 a été démontrée au cours du siècle dernier suivant les travaux de Kolmogorov. En particulier, il est possible de s'assurer que $W(t, \omega)$ soit continu en t pour tout $\omega \in \Omega$, quitte à modifier le processus original sur un ensemble négligeable. Pour les besoins de la cause, un mouvement brownien pourra d'ailleurs être perçu comme étant une marche aléatoire symétrique renormalisée où les étapes se succèdent infiniment rapidement et où le mouvement issu de chacune de ces étapes est infiniment petit. La terminologie « un mouvement brownien » est souvent utilisée dans la littérature pour désigner à la fois une réalisation d'un tel processus ainsi que le processus en question.

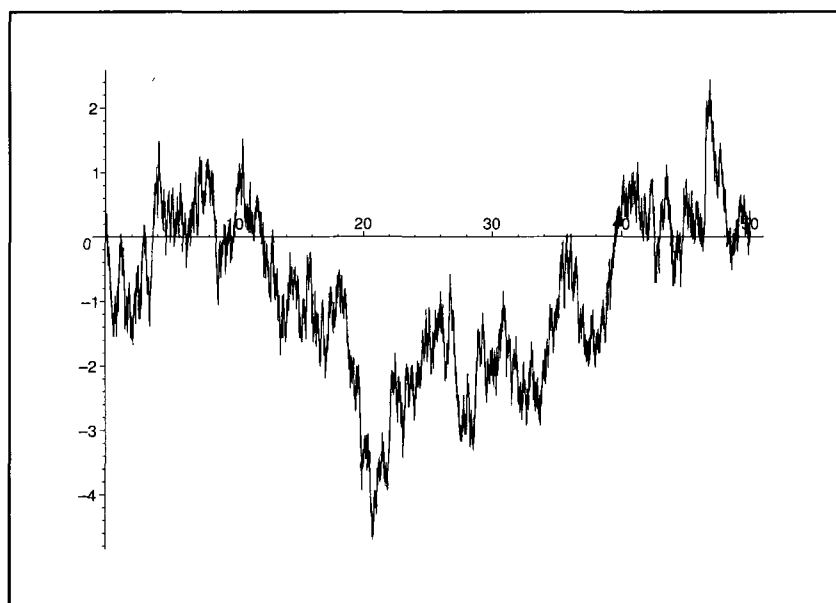


Figure 1.3 : Réalisation d'un mouvement brownien

La figure 1.3 est une approximation (d'une réalisation) d'un mouvement brownien obtenue à l'aide d'une marche symétrique renormalisée en prenant $n = 2000$.

Puisqu'il est possible de le percevoir comme étant la limite d'une marche aléatoire symétrique renormalisée, nous constatons que les « sommets » présents dans la figure 1.2 devraient survenir infiniment souvent (et à intervalles infiniment petits) dans le graphe du mouvement brownien. En effet, le mouvement brownien est un processus continu qui est dérivable nulle

part, une particularité qui nous forcera à recourir au calcul d'Itô.

1.2 Martingale

Définition 1.2.1

Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, T un nombre réel positif et $\{\mathcal{F}(t)\}_{0 \leq t \leq T}$ une filtration de sous- σ -algèbres de \mathcal{F} . Un processus stochastique adapté $\{M(t)\}_{0 \leq t \leq T}$ est une **martingale** si $\mathbb{E}[M(t) | \mathcal{F}(s)] = M(s)$ pour tous $0 \leq s \leq t \leq T$.

En d'autres termes, une martingale est un processus stochastique dont l'espérance conditionnelle au temps t sachant la valeur du processus jusqu'au temps $s < t$ n'est autre que la valeur de ce processus au temps s .

Remarque 1.2.2 - Du point de vue de l'espérance, nous pourrions nous attendre à ce que la valeur de la martingale demeure inchangée à un temps précis dans le futur : il faut néanmoins faire attention de ne pas commettre « l'erreur du parieur ». Considérons à titre d'exemple une **marche aléatoire symétrique** M_k (voir définition 1.1.1) : il s'agit d'une **martingale (discrète)**. En effet, pour chaque lancer d'une pièce de monnaie équilibrée, nous avons $\mathbb{P}(\text{Pile}) = \frac{1}{2} = \mathbb{P}(\text{Face})$, d'où $\mathbb{E}(M_k) = 0$ pour tout k . Supposons par ailleurs qu'une réalisation de $\{M_k\}$ soit telle que $M_4 = 4$. Bien que nous nous attendions en moyenne à obtenir autant de fois « pile » que « face », il est erroné de penser qu'il soit plus probable d'obtenir $M_5 = 3$ plutôt que $M_5 = 5$. Ce que nous dit la propriété de martingale est que $\mathbb{E}(M_5 | M_4 = 4) = M_4 = 4$, ou encore que $\mathbb{P}(M_5 = 3) = \frac{1}{2} = \mathbb{P}(M_5 = 5)$, soit que le fait d'avoir déjà obtenu quatre « pile » consécutifs n'a aucun impact sur les probabilités rattachées au prochain lancer.

La remarque précédente recèle une particularité importante des martingales qui sera utilisée afin de justifier une hypothèse du marché boursier au sein de plusieurs modèles : sachant comment le marché s'est comporté jusqu'à présent, il nous est impossible de savoir si celui-

ci croîtra ou décroîtra dans un avenir même rapproché. Cette dernière affirmation peut être ou ne pas être acceptée comme étant un axiome lors de l'établissement d'un modèle et il serait sage de considérer les implications de ces deux possibilités, surtout lorsque nous tentons de modéliser le marché à un moment précis où nous supposons une tendance haussière. Nous discuterons davantage de cette hypothèse (efficience du marché) au chapitre 2.

Nous avons le résultat important suivant [53] :

Proposition 1.2.3

Un mouvement brownien $W(t)$, pour $t \geq 0$, est une martingale.

Remarque 1.2.4 - La notation abusive $W(t)$ est souvent préférée à $\{W(t)\}_{t \geq 0}$ ou encore $\{W(t, \omega)\}_{t \geq 0}$ dans la littérature pour désigner un mouvement brownien.

1.3 Intégrale de Lebesgue

À la manière de l'intégrale de Riemann qui est définie à l'aide de la limite d'une somme sur une partition, définissons l'intégrale de Lebesgue. Cette section est présentée à titre de rappel pour le lecteur seulement et n'est en aucune façon rigoureuse ou exhaustive.

Idée Dans le cas d'une intégrale de Riemann dans le plan réel, nous considérons le graphe d'une fonction $f(x)$ où $f : \mathbb{R} \rightarrow \mathbb{R}$. Ainsi, il est naturel de partitionner l'axe des abscisses afin de déterminer « l'aire sous la courbe. » Dans le cas d'une intégrale de Lebesgue, nous considérons le graphe d'une fonction mesurable $f(\omega)$ où $f : \Omega \rightarrow \mathbb{R}$. Or, il n'y a pas de façon naturelle de partitionner Ω afin d'effectuer la même tâche. Pour cette raison, nous partitionnerons plutôt l'axe des ordonnées.

Soient donc $f : \Omega \rightarrow \mathbb{R}^+$ une fonction mesurable sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et $\Pi = \{y_0, y_1, y_2, \dots\}$ une partition de \mathbb{R}^+ où $y_0 = 0$. Pour chaque sous-intervalle $[y_i, y_{i+1}[$

de l'image de f , posons

$$A_i = \{\omega \in \Omega \mid y_i \leq f(\omega) < y_{i+1}\}.$$

La figure 1.4 illustre un tel partitionnement.

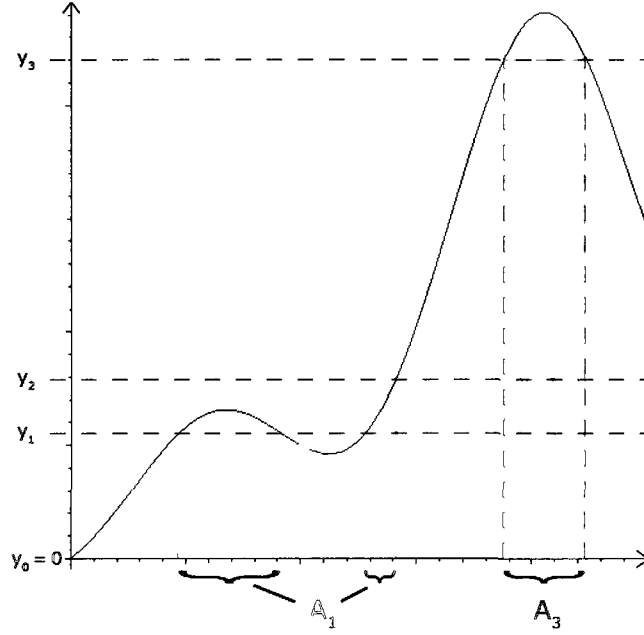


Figure 1.4 : Intégrale de Lebesgue — Partitionnement

Définissons la somme inférieure de Lebesgue comme étant

$$LS_{\Pi}^{-}(f) = \sum_{i=1}^{\infty} y_i \mathbb{P}(A_i)$$

Soit $\|\Pi\|$ la distance maximale entre deux points consécutifs de la partition Π . Nous définissons alors l'intégrale de Lebesgue comme étant la limite de la somme LS_{Π}^{-} lorsque $\|\Pi\| \rightarrow 0$, c'est-à-dire

$$\int_{\Omega} f(\omega) d\mathbb{P}(\omega) = \lim_{\|\Pi\| \rightarrow 0} LS_{\Pi}^{-}(f).$$

Le partitionnement sur l'image de f plutôt que sur son domaine confère à l'intégrale de Lebesgue l'essentiel de sa puissance en contournant complètement la problématique issue d'une fonction admettant une certaine quantité — au plus dénombrable — de discontinuités.

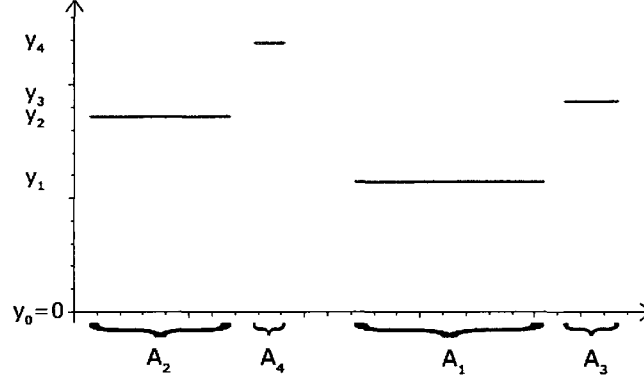


Figure 1.5 : Intégrale de Lebesgue — Discontinuités

Intégrale d'une fonction indicatrice D'abord, considérons le cas simple d'une fonction indicatrice $\mathbf{1}_B$ où $B \in \mathcal{F}$. Alors

$$LS_{\Pi}^{-}(\mathbf{1}_B) = 1 \cdot \mathbb{P}(B) = \mathbb{P}(B)$$

et donc

$$\int_{\Omega} \mathbf{1}_B d\mathbb{P} = \lim_{\|\Pi\| \rightarrow 0} \mathbb{P}(B) = \mathbb{P}(B)$$

Intégrale d'une fonction étagée positive Considérons maintenant le cas d'une fonction mesurable étagée positive E , c'est-à-dire une fonction de la forme $E = \sum_{i \in I} a_i \mathbf{1}_{B_i}$ où les B_i sont des sous-ensembles mesurables, I est de cardinal fini et $a_i \geq 0$ pour $i \in I$. Ainsi,

$$\int_{\Omega} E d\mathbb{P} = \int_{\Omega} \left(\sum_i a_i \mathbf{1}_{B_i} d\mathbb{P} \right) = \sum_i a_i \left(\int_{\Omega} \mathbf{1}_{B_i} d\mathbb{P} \right) = \sum_i a_i \mathbb{P}(B_i)$$

Intégrale d'une fonction mesurable positive Soit f une fonction mesurable positive. Il suffit d'approcher f « par en-dessous » à l'aide d'une fonction mesurable étagée positive. En particulier, nous posons

$$\int_{\Omega} f d\mathbb{P} = \sup \left\{ \int_{\Omega} E d\mathbb{P} \mid 0 \leq E \leq f \text{ et } E \text{ est étagée positive} \right\}.$$

Intégrale d'une fonction mesurable Soit f une fonction mesurable quelconque. Posons

$$f^+(\omega) = \max\{f(\omega), 0\} \quad \text{et} \quad f^-(\omega) = \max\{-f(\omega), 0\},$$

les parties positives et négatives de f respectivement, de telle sorte que $f = f^+ - f^-$ et $|f| = f^+ + f^-$. Alors

$$\int_{\Omega} f \, d\mathbb{P} = \int_{\Omega} f^+ \, d\mathbb{P} - \int_{\Omega} f^- \, d\mathbb{P}.$$

Dans le cas important où l'espace est défini sur les nombres réels, c'est-à-dire $\Omega = \mathbb{R}$, nous avons des outils additionnels pour calculer les intégrales de Lebesgue.

Définition 1.3.1

La plus petite σ -algèbre contenant tous les intervalles fermés $[a, b] \subseteq \mathbb{R}$ est appelée **σ -algèbre de Borel** et notée $\mathcal{B}(\mathbb{R})$. Un élément $B \in \mathcal{B}(\mathbb{R})$ est appelé un **borélien**.

Définition 1.3.2

La **mesure de Lebesgue** sur \mathbb{R} , notée \mathcal{L} , est l'application

$\mathcal{L} : \mathcal{B}(\mathbb{R}) \rightarrow \bar{\mathbb{R}}^+$, où $\bar{\mathbb{R}}^+ := [0, \infty[\cup \{\infty\}$, telle que :

1. $\mathcal{L}([a, b]) = b - a$ pour $a \leq b$;
2. si B_1, B_2, B_3, \dots est une suite de boréliens disjoints deux à deux, alors

$$\mathcal{L}\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mathcal{L}(B_n).$$

Nous avons les deux résultats importants suivants :

Théorème 1.3.3

Soient f une fonction bornée sur \mathbb{R} ainsi que a et b des nombres réels tels que $a < b$. Si l'intégrale de Riemann $\int_a^b f(x) \, dx$ est définie, alors f est Borel-mesurable et les intégrales de Riemann et de Lebesgue sont en accord, c'est-à-dire

$$\int_{[a,b]} f(x) \, d\mathcal{L}(x) = \int_a^b f(x) \, dx$$

Théorème 1.3.4

Soit $F(x)$ la fonction de répartition d'une variable aléatoire $X : \mathbb{R} \rightarrow \mathbb{R}$ admettant la densité de probabilité $f(x)$ et soit $g(x)$ une fonction telle que $\mathbb{E}(|g(X)|)$ soit finie. Alors :

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) dF(x).$$

De façon équivalente, nous avons :

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx.$$

Exemple 1.3.5

Soit X une variable aléatoire suivant une loi normale de moyenne μ et de variance σ^2 . Si $\Phi(x)$ est la fonction de répartition de X , alors $\psi(x) = \Phi'(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ est sa densité de probabilité et donc

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x d\Phi(x) = \int_{-\infty}^{\infty} x\psi(x) dx = \mu.$$

1.4 Calcul d'Itô

1.4.1 Intégrale d'Itô

Nous cherchons à donner un sens à l'expression

$$\int_0^T \Delta(t) dW(t) \tag{1.7}$$

où $W(t)$ est un processus de Wiener (réf. définition 1.1.16) et $\Delta(t)$ est un processus stochastique adapté à la filtration de $W(t)$ pour $t \geq 0$. Si $W(t)$ était dérivable, il suffirait d'appliquer le théorème 1.3.4 afin de calculer la valeur de l'intégrale ci-dessus. Or, tel que mentionné précédemment, malgré le fait que le processus de Wiener soit continu partout, il est dérivable nulle part. Par ailleurs, nous pouvons exploiter le fait qu'il soit continu pour développer un outil à la manière de l'intégrale de Riemann (ou plus précisément de Riemann-Stieltjes). L'outil que nous utiliserons pour donner un sens à l'expression ci-dessus, est l'intégrale d'Itô.

Idée Le concept « d'aire sous la courbe » utilisé dans le cas d'une intégrale de Riemann consiste à estimer l'aire de régions rectangulaires et de prendre la limite d'une somme de telles régions lorsque la base de chaque rectangle tend vers zéro. Ainsi, pour un rectangle borné par x_i et x_{i+1} sur l'axe des abscisses et par 0 et $y_i = \min\{f(x) \mid x_i \leq x \leq x_{i+1}\}$ sur l'axe des ordonnées, nous calculons l'aire¹

$$\Delta(y)\Delta(x) = (y_i - 0)(x_{i+1} - x_i) = y_i(x_{i+1} - x_i).$$

Ceci est dû au fait que nous cherchions à intégrer par rapport à dx . Ainsi, nous cherchons à observer la variation de notre intégrande y_i en fonction du changement d'une unité de x . Et si notre unité était $W(x)$ plutôt que x ? La base de la région rectangulaire en question serait maintenant bornée par $W(x_i)$ et $W(x_{i+1})$. Ainsi, l'aire approximative de cette portion « du graphe » serait

$$\Delta(y)\Delta(W(x)) = (y_i - 0)(W(x_{i+1}) - W(x_i)) = y_i(W(x_{i+1}) - W(x_i)).$$

Somme partielle Soit $\Pi = \{t_0, t_1, \dots, t_n\}$ une partition de $[0, T]$ et soit $\Delta(t)$ le processus stochastique adapté apparaissant dans l'équation (1.7). Calculons une approximation de l'intégrale — à la manière de l'intégrale de Riemann — en calculant la contribution d'une région rectangulaire, c'est-à-dire en supposant que $\Delta(t)$ est constante — de valeur $\Delta(t_j)$ — sur chaque intervalle $[t_j, t_{j+1}[$. Posons $I(t)$ l'approximation de la portion de l'intégrale comprise entre $t_0 = 0$ et $t \leq T$. Si t est compris entre t_0 et t_1 , alors

$$I(t) = \Delta(t_0)[W(t) - W(t_0)] = \Delta(0)W(t).$$

Si par ailleurs t est compris entre t_1 et t_2 , nous avons

$$\begin{aligned} I(t) &= \Delta(t_0)[W(t_1) - W(t_0)] + \Delta(t_1)[W(t) - W(t_1)] \\ &= \Delta(0)W(t_1) + \Delta(t_1)[W(t) - W(t_1)]. \end{aligned}$$

Pour t entre t_2 et t_3 , nous obtenons

$$\begin{aligned} I(t) &= \Delta(t_0)[W(t_1) - W(t_0)] + \Delta(t_1)[W(t_2) - W(t_1)] + \Delta(t_2)[W(t) - W(t_2)] \\ &= \Delta(0)W(t_1) + \Delta(t_1)[W(t_2) - W(t_1)] + \Delta(t_2)[W(t) - W(t_2)]. \end{aligned}$$

¹Notons ici que $\Delta(x)$ représente une variation de la valeur de x et ne fait pas référence au processus stochastique adapté $\Delta(t)$ présent dans l'équation (1.7).

En général, si t est compris entre t_i et t_{i+1} alors

$$I(t) = \sum_{j=0}^{i-1} \Delta(t_j)[W(t_{j+1}) - W(t_j)] + \Delta(t_i)[W(t) - W(t_i)].$$

Limite L'intégrale d'Itô d'un processus stochastique adapté $\Delta(t)$, que nous noterons $I(t) = \int_0^t \Delta(u) dW(u)$, est obtenue en prenant la limite lorsque $\|\Pi\| \rightarrow 0$, c'est-à-dire

$$I(t) = \lim_{\|\Pi\| \rightarrow 0} \sum_{t_i, t_{i+1} \in \Pi} \Delta(t_i)[W(t_{i+1}) - W(t_i)]. \quad (1.8)$$

Cette définition revient à dire que nous « approchons » l'intégrande $\Delta(t)$ par un processus stochastique adapté étagé, de façon similaire à l'intégrale de Lebesgue. Dans le cas de l'intégrale d'Itô, cependant, l'approximation est effectuée au point initial de chaque intervalle plutôt qu'au point admettant la plus petite image sur cet intervalle. Dans cette optique, la limite définie par l'équation (1.8) converge en probabilité dans la mesure où la condition de carré intégrabilité est respectée, soit

$$\int_0^T \Delta^2(t) dt < \infty.$$

Proposition 1.4.1

L'intégrale d'Itô $I(t)$ telle que définie par l'équation (1.8) admet les propriétés importantes suivantes [53] :

1. **(Martingale)** $I(t)$ est une martingale.
2. **(Isométrie d'Itô)** $\mathbb{E}[I^2(t)] = \mathbb{E}[\int_0^t \Delta^2(u) du]$.
3. **(Variation quadratique)** $[I, I](t) = \int_0^t \Delta^2(u) du$.

1.4.2 Processus d'Itô

Le calcul d'Itô est développé pour une famille de processus stochastiques que nous appellerons processus d'Itô.

Définition 1.4.2

Soient $W(t)$ un mouvement brownien et $\mathcal{F}(t)$ une filtration associée pour $t \geq 0$. Un **processus d'Itô** est un processus stochastique de la forme

$$X(t) = X(0) + \int_0^t \Delta(u) dW(u) + \int_0^t \Theta(u) du \quad (1.9)$$

où $X(0)$ est non aléatoire et où $\Delta(t)$ et $\Theta(t)$ sont des processus stochastiques adaptés à la filtration $\mathcal{F}(t)$ pour $t \geq 0$.

Nous supposons évidemment que $\Delta(t)$ est de carré intégrable et que $\int_0^t |\Theta(u)| du < \infty$ afin que $X(t)$ soit défini pour tout $t \geq 0$.

Notons également qu'un processus d'Itô peut être spécifié sous sa forme différentielle, notamment

$$dX(t) = \Delta(t) dW(t) + \Theta(t) dt. \quad (1.10)$$

Proposition 1.4.3

Soit $X(t)$ un processus d'Itô tel que défini par l'équation (1.9). Alors

$$[X, X](t) = \int_0^t \Delta^2(u) du.$$

Nous pouvons ainsi définir l'intégrale par rapport à un processus d'Itô.

Définition 1.4.4

Soit $X(t)$ un processus d'Itô et soit $\Gamma(t)$ un processus stochastique adapté pour $t \geq 0$. L'intégrale de $\Gamma(t)$ par rapport au processus d'Itô $X(t)$ est définie comme étant

$$\int_0^t \Gamma(u) dX(u) = \int_0^t \Gamma(u) \Delta(u) dW(u) + \int_0^t \Gamma(u) \Theta(u) du.$$

1.4.3 Formule d'Itô-Doebelin

Puisque nous avons déterminé une forme d'intégration pour les processus d'Itô, il est naturel de chercher à établir une forme de dérivation. Ainsi, nous pourrions véritablement parler de

« calcul d'Itô. » Étant donnée une fonction f quelconque, nous voudrions avoir une formule pour « dériver » des expressions de la forme $f(X(t))$ où $X(t)$ est un processus d'Itô.

Mouvement brownien

Restreignons-nous d'abord à un mouvement brownien. Soit donc $W(t)$ un mouvement brownien. Si ce processus était dérivable au sens classique, il suffirait d'appliquer la règle de dérivation en chaîne à $f(W(t))$. Or, ce n'est pas le cas. Il faut donc recourir aux résultats importants suivants (voir [53]) :

Théorème 1.4.5 (Formule d'Itô-Doeblin — Mouvement brownien.)

Soit $f(t, x)$ une fonction pour laquelle les dérivées partielles $f_t(t, x)$, $f_x(t, x)$ et $f_{xx}(t, x)$ sont définies et continues, et soit $W(t)$ un processus de Wiener. Alors, pour tout $T \geq 0$, nous avons

$$\begin{aligned} f(T, W(T)) = f(0, W(0)) &+ \int_0^T f_t(t, W(t)) dt \\ &+ \int_0^T f_x(t, W(t)) dW(t) + \frac{1}{2} \int_0^T f_{xx}(t, W(t)) dt. \end{aligned} \quad (1.11)$$

Ainsi, en exprimant ce résultat sous une forme différentielle, nous obtenons :

Lemme 1.4.6 (Lemme d'Itô — Mouvement brownien.)

Sous les mêmes conditions que le théorème 1.4.5, nous avons

$$df(t, W(t)) = f_t(t, W(t)) dt + f_x(t, W(t)) dW(t) + \frac{1}{2} f_{xx}(t, W(t)) dt. \quad (1.12)$$

Dans l'élaboration du résultat précédent, nous en déduisons quelques formules importantes, notamment

$$dW(t) dW(t) = dt, \quad dt dW(t) = dW(t) dt = 0, \quad dt dt = 0. \quad (1.13)$$

Nous réutiliserons ces formules régulièrement en manipulant des équations (stochastiques) aux dérivées partielles.

Processus d'Itô

De façon analogue, nous avons les résultats suivants pour un processus d'Itô $X(t)$ quelconque :

Théorème 1.4.7 (Formule d'Itô-Doeblin — Processus d'Itô)

Soit $f(t, x)$ une fonction pour laquelle les dérivées partielles $f_t(t, x)$, $f_x(t, x)$ et $f_{xx}(t, x)$ sont définies et continues, et soit $X(t)$ un processus d'Itô, pour $t \geq 0$, tel que représenté par l'équation (1.9). Alors, pour tout $T \geq 0$, nous avons

$$\begin{aligned} f(T, X(T)) = f(0, X(0)) &+ \int_0^T f_t(t, X(t)) dt + \int_0^T f_x(t, X(t)) \Delta(t) dW(t) \\ &+ \int_0^T f_x(t, X(t)) \Theta(t) dt + \frac{1}{2} \int_0^T f_{xx}(t, X(t)) \Delta(t)^2 dt. \end{aligned} \quad (1.14)$$

Or, si le théorème 1.4.7 est utile pour déterminer explicitement la valeur de $f(t, X(t))$ et peut simplifier le calcul de certaines intégrales d'Itô, le résultat suivant — mieux connu sous le nom de « lemme d'Itô » — s'avère quant à lui essentiel dans l'élaboration de l'équation de Black-Scholes et nous donne une formule de « dérivation » pour un processus d'Itô. Il s'agit notamment de l'équivalent du théorème ci-haut mentionné exprimé sous une forme différentielle.

Lemme 1.4.8 (Lemme d'Itô)

Sous les mêmes conditions que le théorème 1.4.7, nous avons

$$\begin{aligned} df(t, X(t)) = f_t(t, X(t)) dt &+ f_x(t, X(t)) \Delta(t) dW(t) \\ &+ f_x(t, X(t)) \Theta(t) dt + \frac{1}{2} f_{xx}(t, X(t)) \Delta^2(t) dt. \end{aligned} \quad (1.15)$$

À l'aide de l'équation (1.10), nous pouvons réécrire l'équation ci-dessus sous la forme

$$df(t, X(t)) = f_t(t, X(t)) dt + f_x(t, X(t)) dX + \frac{1}{2} f_{xx}(t, X(t)) \Delta^2(t) dt. \quad (1.16)$$

1.5 Black-Scholes

Remarque 1.5.1 - Les termes « investisseur », « souscripteur », « acheteur » et « vendeur » sont employés sous leur forme masculine dans cette section pour désigner sans discrimination une personne pouvant être une femme ou un homme dans l'unique but d'alléger le texte.

Étant donnée la diversité et l'étendue des marchés financiers à travers le monde, il y a plusieurs options qui s'offrent aux individus concernant les méthodes d'investissement. En particulier, il y a la possibilité d'établir des ententes contractuelles entre deux parties permettant ainsi à l'acheteur (respectivement au vendeur) de se prévaloir (respectivement de se départir) d'un certain nombre d'actions à un temps ultérieur et à un prix préalablement fixés au moment de l'entente.

Définition 1.5.2

Un **produit dérivé** est une entente contractuelle dont la valeur (fictive) fluctue en fonction de celle d'un actif sous-jacent, réel ou théorique.

Définition 1.5.3

Une **option européenne** est un produit dérivé qui donne le droit (et non l'obligation) au souscripteur d'acheter ou de vendre un actif à un prix préalablement fixé (appelé **strike**, ou prix d'exercice) à une date prédéterminée (appelée date de maturité). Une option d'achat (respectivement de vente) est communément appelée un **call** (respectivement un **put**).

L'équation de Black-Scholes est utilisée afin de modéliser l'évolution du prix d'une option d'achat de style européen. L'idée consiste d'abord à modéliser l'évolution d'une action quelconque — dont le prix est dicté par un processus stochastique — sur un marché boursier et ensuite à établir le prix d'une option de type call dont l'actif sous-jacent est le prix de l'action tel que dicté par ce modèle. Ces résultats, établis par Merton [44] et basés sur les ouvrages de Black et Scholes, auront valu à Merton et à Scholes le prix de la Banque de Suède en sciences

économiques en mémoire d'Alfred Nobel, plus couramment appelé prix Nobel d'économie². Ce prix leur aura été octroyé en 1997, soit deux ans après la mort de Black qui n'aura été mentionné qu'à titre de contributeur³.

1.5.1 Modèle de Black-Scholes

Considérons un portefeuille dont la valeur au temps t est $X(t)$ et qui investit à la fois dans un marché rapportant un taux d'intérêt constant r et dans une action dont la valeur $S(t)$ est modélisée par un mouvement brownien géométrique tel que dicté par l'équation

$$dS(t) = \alpha S(t) dt + \sigma S(t) dW(t) \quad (1.17)$$

où α et σ sont des nombres réels fixés et $W(t)$ est un mouvement brownien.

Soit $\Delta(t)$ le nombre de parts d'une action détenues par l'investisseur au temps t . Ainsi, la portion du portefeuille consacrée à l'action est $\Delta(t)S(t)$. La valeur restante du portefeuille, soit $X(t) - \Delta(t)S(t)$, est investie dans le marché à intérêt constant. La valeur du portefeuille variant uniquement en fonction de ces deux facteurs, nous avons :

$$\begin{aligned} dX(t) &= \Delta(t) dS(t) + r[X(t) - \Delta(t)S(t)] dt \\ &= rX(t) dt + \Delta(t)(\alpha - r)S(t) dt + \Delta(t)\sigma S(t) dW(t) \end{aligned} \quad (1.18)$$

après substitution de la valeur de $dS(t)$.

Soit $K > 0$ le prix d'exercice convenu préalablement, c'est-à-dire le prix auquel le souscripteur a le droit (et non l'obligation) d'acheter un certain nombre de parts de l'action — ce nombre étant $\Delta(t)$ à tout moment $t > 0$ — et soit $T > 0$ la date de maturité de l'option. Si la valeur $S(t)$ des parts à un temps $0 < t \leq T$ est supérieure à K , alors une transaction serait avantageuse pour l'acheteur et celui-ci bénéficierait d'un coût d'achat inférieur à la valeur actuelle du marché ; sinon, la transaction serait désavantageuse et ne serait donc pas effectuée.

²Il s'agit du seul prix géré par la Fondation Nobel qui n'a pas été créé par le testament d'Alfred Nobel, d'où la confusion quant à son appellation.

³La fondation Nobel n'octroie aucun prix à titre posthume.

Rappelons toutefois qu'une transaction ne peut être effectuée qu'au temps T pour une option de style européen. Ainsi, l'option d'achat offre au souscripteur une économie de $(S(T) - K)^+$ au temps $T > 0$, où

$$(S(T) - K)^+ = \max(S(T) - K, 0).$$

Nous cherchons à établir un « prix raisonnable » (en fait, un prix idéal) pour l'option. Le modèle de Black-Scholes émet l'hypothèse que celui-ci ne devrait varier qu'en fonction du temps et du prix de l'action à ce temps. Notons donc par $c(t, x)$ le prix recherché de l'option au temps t si la valeur du stock à ce temps est $S(t) = x$.

Supposons que nous investissons dans l'action S au temps $t > 0$. Nous investirons donc un montant équivalent à $S(t)$. Puisque l'action est dotée d'un taux d'intérêt fixe r composé continûment, la valeur de notre investissement serait $S(T) = e^{r(T-t)}S(t)$ à tout temps $T > t$ dans le futur. Il s'agit de la valeur *capitalisée* de l'action. Inversement, sachant la valeur $S(T)$ d'un investissement dans le futur, nous pouvons calculer la valeur $S(t) = e^{-r(T-t)}S(T)$ que nous aurions eu à investir à tout moment $t < T$. Il s'agit de la valeur *actualisée* de l'action.

Considérons les prix actualisés de l'action et du portefeuille sous étude, soient $e^{-rt}S(t)$ et $e^{-rt}X(t)$ respectivement. Notons que, dans ce cas, l'actualisation est effectuée par rapport au moment initial ($t = 0$) sachant les valeurs au temps futur $t > 0$.

Notre portefeuille a un capital initial $X(0)$ et nous cherchons à ce que la valeur $X(t)$ du portefeuille soit en accord avec la valeur $c(t, S(t))$ de l'action pour tout $t \in [0, T[$. Ceci est le cas si et seulement si

$$e^{-rt}X(t) = e^{-rt}c(t, S(t)) \quad \text{pour tout } t \in [0, T[. \quad (1.19)$$

Il suffit donc de considérer le cas où

$$\begin{cases} X(0) = c(0, S(0)) \\ d(e^{-rt}X(t)) = d(e^{-rt}c(t, S(t))) \end{cases} \quad \text{pour tout } t \in [0, T[. \quad (1.20)$$

$$(1.21)$$

En effet, en intégrant (1.21) sur $[0, t[$, nous obtenons

$$e^{-rt}X(t) - X(0) = e^{-rt}c(t, S(t)) - c(0, S(0))$$

ce qui en vertu de l'équation (1.20) revient à étudier la situation souhaitée telle qu'énoncée par l'équation (1.19).

Posons $f(t, x) = e^{-rt}x$ et appliquons le lemme d'Itô. Nous obtenons

$$\begin{aligned} d(e^{-rt}S(t)) &= df(t, S(t)) \\ &= f_t(t, S(t)) dt + f_x(t, S(t)) dS(t) + \frac{1}{2}f_{xx}(t, S(t)) dS(t)dS(t) \\ &= (\alpha - r)e^{-rt}S(t) dt + \sigma e^{-rt}S(t) dW(t) \end{aligned} \quad (1.22)$$

et

$$\begin{aligned} d(e^{-rt}X(t)) &= df(t, X(t)) \\ &= f_t(t, X(t)) dt + f_x(t, X(t)) dX(t) + \frac{1}{2}f_{xx}(t, X(t)) dX(t)dX(t) \\ &= \Delta(t)[(\alpha - r)e^{-rt}S(t) dt + \sigma e^{-rt}S(t) dW(t)]. \end{aligned} \quad (1.23)$$

Remarquons des équations (1.22) et (1.23) que $d(e^{-rt}X(t)) = \Delta(t) d(e^{-rt}S(t))$.

En vertu du lemme d'Itô, nous avons :

$$dc(t, S(t)) = c_t(t, S(t)) dt + c_x(t, S(t)) dS(t) + \frac{1}{2}c_{xx}(t, S(t)) dS(t)dS(t). \quad (1.24)$$

Or, les équations (1.13) et (1.17) donnent

$$\begin{aligned} dS(t)dS(t) &= (\alpha S(t) dt + \sigma S(t) dW(t))^2 \\ &= \alpha^2 S^2(t) dt dt + 2\alpha\sigma S^2(t) dt dW(t) + \sigma^2 S^2(t) dW(t)dW(t) \\ &= \sigma^2 S^2(t) dt. \end{aligned} \quad (1.25)$$

Ainsi, après substitution des équations (1.17) et (1.25) dans (1.24), nous obtenons

$$\begin{aligned} dc(t, S(t)) &= \left[c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t)c_{xx}(t, S(t)) \right] dt \\ &\quad + \sigma S(t)c_x(t, S(t)) dW(t). \end{aligned} \quad (1.26)$$

Posons $f(t, x) = e^{-rt}x$. En vertu du lemme d'Itô, le prix actualisé de $c(t, S(t))$ est donné par :

$$\begin{aligned}
d(e^{-rt}c(t, S(t))) &= df(t, c(t, S(t))) \\
&= f_t(t, c(t, S(t))) dt + f_x(t, c(t, S(t))) dc(t, S(t)) \\
&\quad + \frac{1}{2}f_{xx}(t, c(t, S(t))) dc(t, S(t))dc(t, S(t)) \\
&= -re^{-rt}c(t, S(t)) dt + e^{-rt} dc(t, S(t)) \\
&= e^{-rt} \left[-rc(t, S(t)) + c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) \right. \\
&\quad \left. + \frac{1}{2}\sigma^2 S^2(t)c_{xx}(t, S(t)) \right] dt + e^{-rt}\sigma S(t)c_x(t, S(t)) dW(t) \quad (1.27)
\end{aligned}$$

après substitution de la valeur de $dc(t, S(t))$ telle que donnée par l'équation (1.26).

Ainsi, en vertu des équations (1.22) et (1.27), nous en concluons que la condition

$d(e^{-rt}X(t)) = d(e^{-rt}c(t, S(t)))$ est vérifiée si et seulement si

$$\begin{aligned}
&\Delta(t)(\alpha - r)S(t) dt + \Delta(t)\sigma S(t) dW(t) \\
&= \left[-rc(t, S(t)) + c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t)c_{xx}(t, S(t)) \right] dt \\
&\quad + \sigma S(t)c_x(t, S(t)) dW(t). \quad (1.28)
\end{aligned}$$

Cette égalité tient si et seulement si pour tout $t \in [0, T[$ nous avons

$$1) \Delta(t) = c_x(t, S(t))$$

$$\begin{aligned}
2) (\alpha - r)S(t)c_x(t, S(t)) = \\
-rc(t, S(t)) + c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t)c_{xx}(t, S(t))
\end{aligned}$$

c'est-à-dire si et seulement si

$$\left\{ \begin{array}{l} \Delta(t) = c_x(t, S(t)) \end{array} \right. \quad (1.29)$$

$$\left\{ \begin{array}{l} rc(t, S(t)) = c_t(t, S(t)) + rS(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t)c_{xx}(t, S(t)). \end{array} \right. \quad (1.30)$$

Les équations (1.29) et (1.30) assurent que la condition établie par l'équation (1.19) est respectée. Ainsi, $e^{-rt}X(t) = e^{-rt}c(t, S(t))$ et donc $X(t) = c(t, S(t))$ pour tout $t \in [0, T[$. Lorsque

$t \rightarrow T$ nous avons $X(T) = c(T, S(T)) = (S(T) - K)^+$, d'où la condition terminale $c(T, x) = (x - K)^+$.

Nous cherchons donc une fonction continue $c(t, x)$ qui satisfasse à l'équation aux dérivées partielles

$$c_t(t, x) + rxc_x(t, x) + \frac{1}{2}\sigma^2 x^2 c_{xx}(t, x) = rc(t, x) \quad (1.31)$$

pour tout $t \in [0, T[$ ainsi que la condition terminale

$$c(T, x) = (x - K)^+. \quad (1.32)$$

1.5.2 Formule de Black-Scholes

Il est question ici d'établir une solution à l'équation (1.31) avec la condition terminale définie par l'équation (1.32). Afin de résoudre une telle équation aux dérivées partielles, il faut des conditions sur les frontières, notamment lorsque $x = 0$ et $x = \infty$. Lorsque $x = 0$, l'équation (1.31) nous donne $c_t(t, 0) = rc(t, 0)$, d'où $c(t, 0) = e^{rt}c(0, 0)$. En substituant $t = T$ dans cette dernière égalité et en sachant de l'équation (1.32) que $c(T, 0) = (0 - K)^+ = 0$, nous obtenons $0 = e^{rT}c(0, 0)$, d'où $c(0, 0) = 0$. Ainsi,

$$c(t, 0) = e^{rt}c(0, 0) = 0 \quad \text{pour tout } t \in [0, T]. \quad (1.33)$$

Lorsque $x = \infty$, la situation est légèrement plus compliquée : comme $c(t, x)$ n'est pas bornée, il faut plutôt se résigner à établir le comportement souhaité de la fonction lorsque $x \rightarrow \infty$. Rappelons que $x = S(t)$ est le prix de l'action au temps t . L'économie effectuée sur la transaction (en supposant bien sûr qu'elle soit avantageuse, sans quoi elle n'est pas effectuée) au temps T est $S(T) - K = x - K$. Si la transaction était effectuée à un temps $t < T$, alors l'économie effectuée serait $S(t) - e^{-r(T-t)}K = x - e^{-r(T-t)}K$ — il ne s'agit ici en fait que d'actualiser la valeur de K au temps $t < T$, le prix de $S(t)$ étant automatiquement ajusté. Ainsi, le prix $c(t, x)$ de notre option au temps t devrait se comporter comme $x - e^{-r(T-t)}K$, ce qui lorsque $x \rightarrow \infty$ nous donne

$$\lim_{x \rightarrow \infty} \left[c(t, x) - \left(x - e^{-r(T-t)}K \right) \right] = 0 \quad \text{pour tout } t \in [0, T[. \quad (1.34)$$

Les équation (1.33) et (1.34) sont les conditions aux frontières recherchées.

La solution à l'équation aux dérivées partielles (1.31) avec les conditions aux frontières données par les équations (1.33) et (1.34) est (voir [53]) :

$$c(t, x) = xN(d_+(T - t, x)) - Ke^{-r(T-t)}N(d_-(T - t, x)), \quad 0 \leq t < T, \quad x > 0 \quad (1.35)$$

où

$$d_{\pm}(\tau, x) = \frac{1}{\sigma\sqrt{\tau}} \left[\log \frac{x}{K} + \left(r \pm \frac{\sigma^2}{2} \right) \tau \right] \quad (1.36)$$

et $N(\cdot)$ est la fonction de répartition de la distribution normale centrée et réduite, notamment

$$N(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{-y}^{\infty} e^{-\frac{z^2}{2}} dz. \quad (1.37)$$

CHAPITRE 2

Modèle classique

Dans le chapitre précédent, nous avons construit le modèle de Black-Scholes et nous avons discuté de certaines de ses propriétés importantes. Mais quels modèles sont réellement utilisés de nos jours ? Quels outils ont déjà été d'usage courant et le sont-ils encore ?

Nous chercherons dans ce chapitre à exposer divers modèles développés et utilisés au cours du dernier siècle et à discuter des problèmes qui se rattachent à leur application dans le domaine de la finance.

2.1 Un peu d'histoire

2.1.1 Brown et le mouvement brownien

En 1827, alors qu'il observait sous le microscope du pollen immergé dans l'eau, Robert Brown constata que de minuscules particules au sein des graines de pollen se mouvaient de façon très rapide et irrégulière. Bien qu'il ne put lui-même expliquer le phénomène, qui avait d'ailleurs déjà été décrit dans une publication [22] de Jan Ingen-Housz datant de 1784, ce dernier ayant plutôt observé des particules de charbon flottant sur une surface d'alcool,

la découverte fut attribuée à Brown et le mouvement brownien a été nommé ainsi en son honneur.

2.1.2 Bachelier, Einstein et la marche aléatoire

Il fallut toutefois attendre la thèse doctorale de Louis Jean-Baptiste Alphonse Bachelier¹, en 1900, pour obtenir un modèle mathématique décrivant le mouvement brownien. Bachelier décida d'établir une méthode pour l'élaboration du prix d'une action, la technique utilisée jusqu'alors s'apparentant à la clairvoyance. Dans sa thèse [3], il modélisa les fluctuations du marché en admettant l'hypothèse que le prix d'une action avait à tout moment, à probabilités égales, la possibilité d'augmenter ou de diminuer. Bachelier venait tout juste d'élaborer la théorie des marches aléatoires et, sans le savoir, offrait un cadre mathématique décrivant tout mouvement brownien.

Dans sa thèse, Bachelier établit la valeur d'une option dont le prix de l'actif sous-jacent est modélisé par un mouvement brownien. Néanmoins, les travaux de Bachelier portant sur l'application du mouvement brownien aux théories de la spéculation et du jeu n'eurent aucun impact direct sur le monde de la finance avant les années 1950, l'élite française des mathématiques (Borel, Lebesgue, Lévy, etc.) lui reprochant un manque de rigueur. Le tempérament explosif de Bachelier et une erreur de jugement de la part de Paul Pierre Lévy donnèrent d'ailleurs lieu à une dispute malheureuse entre les deux mathématiciens qui les accompagna jusqu'à la mort de Bachelier en 1946. Mathématicien prolifique et en avance sur son époque, Bachelier fut victime de son sort et, suite à une évaluation trop ordinaire de sa thèse et à une lettre de Lévy, il eut énormément de difficulté à se trouver un poste dans une institution académique.

Au courant des années 1905 et 1906, Albert Einstein et Marian Smoluchowski firent constater l'importance du mouvement brownien au monde de la physique en conjecturant que celui-

¹Thorvald N. Thiele effleura le sujet dans une publication [62] de 1880 portant sur les moindres carrés, mais celle-ci passa sous silence.

ci pouvait être utilisé pour prouver indirectement l'existence des atomes. C'est Einstein qui proposa de modéliser le mouvement des particules décrites par Brown à l'aide d'une marche aléatoire [15]. Il semblerait toutefois qu'Einstein ne connaissait pas l'existence des travaux de Bachelier à cette époque et que le traitement mathématique du mouvement brownien qu'il fit dans ses travaux de 1905 était beaucoup moins élégant que celui proposé dans la thèse de Bachelier cinq années auparavant.

2.1.3 Wiener et son processus

Alors que Bachelier avait établi la fonction de répartition du processus mathématique décrivant un mouvement brownien dans sa thèse, il fallut attendre jusqu'en 1923 pour que Norbert Wiener en fournisse un modèle mathématique rigoureux et complet [63]. C'est en son honneur que le processus stochastique (continu) sous-jacent à un mouvement brownien quelconque est souvent nommé, en mathématiques, « processus de Wiener ».

2.1.4 De Bachelier à Black-Scholes

Les travaux de Bachelier donnèrent lieu aux travaux d'autres mathématiciens prolifiques au cours du XX^e siècle, notamment ceux de Kolmogorov et d'Itô ; au moment de la mort de Bachelier en 1946, les mathématiciens comprenaient enfin sa véritable contribution. Il fallut néanmoins attendre quelques années supplémentaires afin que les travaux de Bachelier aient des répercussions dans le monde de la finance.

Pendant la première moitié du siècle, ce sont les travaux de gens tels Irving Fisher et John Burr Williams qui dictèrent les modèles mathématiques utilisés pour l'élaboration du prix d'une action. Leurs travaux respectifs datant de 1930 et 1938 proposaient d'évaluer le prix d'une action en considérant tout simplement sa valeur actualisée par rapport à certains facteurs sélectionnés, en particulier le temps. C'est la méthode du « *Discounted cash flow* »

(DCF) »². Tout comme ses prédécesseurs, Williams eut une énorme difficulté à faire valoir ses travaux³ et l'influence que ceux-ci exercèrent sur le monde de la finance fut énorme. Or, en 1950, un jeune doctorant du nom de Harry Max Markowitz, étudiant alors à l'Université de Chicago sous l'influence d'économistes importants tel Leonard Jimmie Savage, considéra que le modèle proposé par Williams n'incorporait pas le concept de « risque ». Markowitz produisit au cours des années qui suivirent ses travaux sur les portefeuilles. Ceux-ci soutenaient que la diversification d'un portefeuille contribue à faire diminuer le risque associé au portefeuille et qu'il est donc primordial de quantifier l'interdépendance des différentes actions en calculant leur corrélation. Or, Markowitz abandonna temporairement l'idée étant donné le nombre de calculs requis, une tâche impensable à l'époque.

Vers le début des années 1950, Paul Anthony Samuelson, professeur d'économie au Massachusetts Institute of Technology (MIT) à ce moment, reçoit une carte postale de Leonard Jimmie Savage lui disant qu'il devait à tout prix lire la thèse de Bachelier, ce qu'il fit. Ce fut une véritable révélation pour Samuelson qui décida désormais de se consacrer entièrement au sujet. Pendant plusieurs années, il travailla à améliorer le cadre mathématique proposé dans la thèse de Bachelier et étudia son application au monde de la finance de cette époque. Le message était lancé : il y avait beaucoup à faire dans le domaine de l'économie et l'ouvrage de Bachelier avait énormément à y contribuer. Les véritables outils économiques ont été développés à partir de cette époque.

Il fallut attendre l'année 1960 afin que l'on donne aux travaux de Markowitz un nouveau souffle : un jeune doctorant, à qui on avait suggéré d'aller rendre visite à Markowitz, vint cogner à sa porte. William Forsyth Sharpe, dont l'actuelle thèse n'aboutissait pas, décida de reprendre les travaux de Markowitz. Son idée était simple : il se demanda ce qui se produirait si tous les individus actifs dans le marché agissaient selon les mêmes règles, notamment celles établies plus tôt par Markowitz. Sharpe en conclut qu'il n'y aurait alors pas un portefeuille efficient pour chaque individu, mais plutôt un seul : le « portefeuille du marché ».

²Traduction approximative : actualisation des flux monétaires (AFM).

³Il dut d'ailleurs assumer une partie des coûts pour la publication de sa thèse.

Ainsi, analyser l'interdépendance entre les diverses actions revient à comparer chaque action « au marché ». Sharpe proposa d'attribuer à chaque action un indice de volatilité — le « beta », ou β de l'action — représentant sa tendance à fluctuer en fonction des variations du marché. À cet effet, on attribue à une action qui fluctue exactement comme le marché global, c'est-à-dire qui augmente et diminue au même rythme, une valeur de $\beta = 1$ alors qu'une action qui diminue de deux unités lorsque le marché augmente d'une unité se verrait octroyer une valeur de $\beta = -2$. On proposa à Sharpe un poste à l'Université de Chicago pour ses travaux, mais le *Journal of Finance* rejetta sa première demande de publication. Publié en 1964 après une demande de changement d'examineur, le modèle de Sharpe [52] atteint une notoriété internationale quelques années après sa publication : connu sous l'acronyme **CAPM** (*Capital Asset Pricing Model*), celui-ci est encore enseigné dans la majorité des institutions académiques. Plusieurs études [45][49] ont cependant recensé une mauvaise utilisation de ce modèle par un bon nombre de personnes qui l'emploient régulièrement dans le monde des affaires. Sharpe venait enfin de permettre une première quantification du risque.

Sous la supervision de Samuelson émergèrent des étudiants qui allaient avoir un impact remarquable dans le monde de l'économie : Robert C. Merton et Myron Samuel Scholes. L'histoire débuta toutefois au MIT avec un dénommé Fischer Sheffey Black. Nouvellement obsédé par un modèle d'établissement de prix des actions développé par Jack L. Treynor quelques années auparavant, mais jamais publié, Black s'intéressait, comme d'innombrables personnes avant lui, à établir une formule pour évaluer des options et des warrants. Pour y arriver, il établit un système d'équations différentielles qu'il n'arrivait toutefois pas à résoudre ; le problème fut mis de côté pour quelque temps, jusqu'au moment où on lui offrit un poste au MIT. C'est alors qu'il rencontra Scholes et que les deux commencèrent à collaborer sur le projet. Scholes jugea utile de discuter de la problématique avec Merton, un autre collègue du MIT, qui s'impliqua partiellement. Leurs travaux furent rassemblés dans un ouvrage qu'ils tentèrent de publier dans le *Journal of Political Economy* en octobre 1970 : le monde de la finance arborant un sens de l'humour très particulier, leur soumission fut rejetée. Il fallut que des économistes de renommée de l'Université de Chicago interviennent — une faveur qu'Eu-

gene Francis Fama et Merton Howard Miller leur accordèrent en tant qu'amis — en exerçant du lobbying auprès des éditeurs du journal afin qu'ils y jettent un second coup d'oeil. L'article [5] fut finalement imprimé en 1973, quelques semaines suivant l'inauguration du *Chicago Board Options Exchange* en avril. C'est ainsi que naquit le modèle de « **Black-Scholes** ».

2.2 Le modèle classique et la réalité

Le modèle de Black-Scholes, également appelé modèle de Black-Scholes-Merton par certains — remarquons d'ailleurs que seuls Scholes et Merton ont pu profiter de la reconnaissance de la société Nobel pour ces travaux —, est encore très largement utilisé de par le monde pour déterminer la valeur de certaines options. C'est le modèle que nous avons développé dans la section 1.5.1 du présent document. Le CAPM, développé par Sharpe, ainsi que les modèles de type ARCH, GARCH, FIGARCH, etc. sont également d'usage courant. Quelles problématiques, donc, se présentent quant à l'utilisation de ces modèles ?

D'abord, comme l'a mentionné Black, *il faut recourir à plusieurs hypothèses réductrices afin de rendre un modèle utilisable et maniable*. Ainsi, tous les modèles discutés dans ce chapitre ont adopté les hypothèses de Bachelier, notamment que *les incréments sont indépendants, stationnaires et normalement distribués*. Ces conséquences découlent de l'adoption de la marche aléatoire (et donc du mouvement brownien) comme modèle sous-jacent à la fluctuation des prix. Or, les données empiriques recueillies au cours des années tendent à montrer que ces hypothèses sont inappropriées pour modéliser les marchés boursiers.

2.2.1 Données aberrantes et queues lourdes

Benoît B. Mandelbrot a consacré la majorité de sa carrière à chercher de l'ordre dans le désordre : en 1975, il publie les travaux [38] qui établiront les fondements des mathématiques fractales, cette branche qui cherche à établir une structure dans le chaos. Or, dès 1951, il s'intéresse à divers domaines dont l'économie. Vers le début des années 1960, alors qu'il

est employé chez IBM, Mandelbrot travaille avec des ordinateurs pour analyser la distribution des revenus dans la société. Il corrobore les résultats de Vilfredo Federico Damaso Pareto en déterminant que cette distribution suit une loi de puissance inverse, ou distribution de Pareto. Alors qu'il visitait l'Université de Harvard afin d'y donner une conférence sur son travail actuel, Mandelbrot croisa un tableau sur lequel se trouvait un graphique quasi identique à ceux que son étude sur la distribution des revenus avait produits. Or, ce graphique représentait l'évolution des prix du coton aux États-Unis. Émerveillé par l'apparition d'une même situation dans deux contextes complètement disjoints en apparence, Mandelbrot se mit à étudier les prix du coton à l'aide des ordinateurs d'IBM.

La coïncidence est parfaite : alors qu'il commence à s'intéresser aux fluctuations du prix d'un produit spécifique, la théorie de Bachelier refait surface dans le monde de l'économie. Qu'il le veuille ou non, Mandelbrot est exposé à l'opinion des experts de tous les coins de la planète qui clament haut et fort la normalité de ces fluctuations. Rapidement, les prix du coton se mettent à parler : il existe des données aberrantes. Beaucoup de données aberrantes.

Mandelbrot et les données aberrantes

Dès le moment où la théorie de Bachelier fut introduite dans le monde de la finance, des chercheurs commencèrent à vérifier les données et y constatèrent des anomalies. La distribution empirique ne colle pas entièrement à la théorie : il y a des données aberrantes vers les deux extrémités qui sont anormalement élevées. Néanmoins, les chercheurs décident d'ignorer ces données et de les retrancher de leurs études.

Il fallut attendre 1962 afin que Mandelbrot fournisse une analyse détaillée, à l'aide d'ordinateurs, des prix du coton aux États-Unis sur une période de plus de 100 ans pour obtenir une vision plus illuminée et quelque peu alarmante de la situation. Non seulement y a-t-il des données anormalement élevées aux extrémités de la distribution empirique, mais elles le sont *beaucoup trop et en trop grand nombre*.

Les résultats étaient clairs et irréfutables. Loin de bien se comporter et d'être normalement distribués comme le suggérait jusqu'alors la théorie classique, les prix du coton bondissaient partout immodérément. Leur variance, plutôt que de demeurer fixe tel qu'attendu, variait par un facteur de cent et ne se stabilisait jamais autour d'une valeur constante.

- *The (mis)Behavior of Markets*, Benoît B. Mandelbrot [42]

Paul H. Cootner republia l'analyse de Mandelbrot dans son livre [13] de 1964 et lança le message aux économistes qu'*il fallait dorénavant porter attention aux observations empiriques inconfortables qu'ils balayaient jusqu'alors sous le tapis*. Suite à l'influence marquée de cet ouvrage, plusieurs économistes ont commencé à porter un nouveau regard sur les données financières.

Aujourd'hui, il est bien connu que les fluctuations de prix des marchés boursiers ne sont pas normalement distribuées. En effet, au cours du dernier siècle, plusieurs événements se sont produits dans le domaine économique avec des probabilités si minimes qu'elles n'auraient jamais dû se produire dans toute l'histoire de notre planète si l'on admet l'hypothèse de normalité. Le lundi, 19 octobre 1987 — journée fatidique communément appelée « Black Monday » —, l'impensable se produisit : contre toute attente, les marchés boursiers à travers le monde s'effondrèrent.

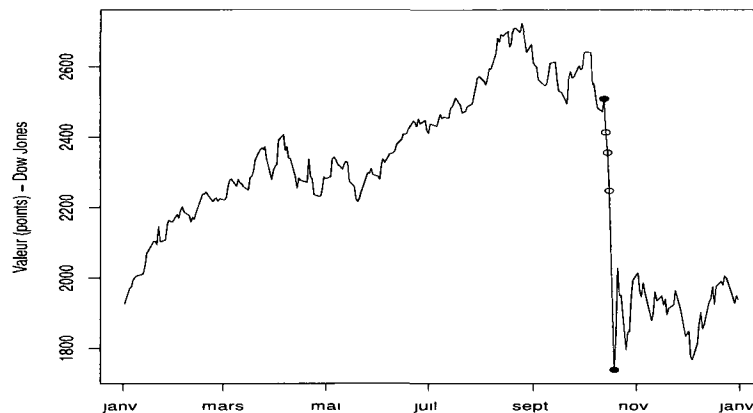


Figure 2.1 : Krach d'octobre 1987 (« Black Monday »)

Le Dow Jones encaisse une perte de 507.99 points sur 2246.73, pour une chute relative de 22.61%. C'est la chute la plus importante de l'histoire et, entre le 13 et le 19 octobre, elle aura fait perdre au Dow Jones 30.68% de sa valeur. En fonction de la fenêtre temporelle utilisée pour l'estimation de la distribution normale des données, la chute du 19 octobre est située environ entre 18 et 22 écarts types de la moyenne. La probabilité qu'un tel événement se produise est inférieure à 10^{-50} .

Observons les données du Dow Jones (DJI) entre le 1^{er} octobre 1928 et le 31 mai 2010. Restreignons-nous seulement aux valeurs quotidiennes de fermeture de l'indice boursier sur cette période ; ainsi, nous ne prenons pas en considération la volatilité du marché au sein d'une journée donnée, mais seulement par rapport aux journées précédente et suivante. Ceci a déjà pour effet d'éliminer des données aberrantes : pensons au 6 mai 2010 où l'indice a perdu 3.15% de sa valeur par rapport à la journée précédente, mais a encaissé une chute de près de 9% au cours de la journée.

Le graphique suivant répartit les variations quotidiennes du DJI dans 200 classes différentes, selon leur taille. Superposé à l'histogramme se trouve la courbe normale qui devrait dicter — en supposant la marche aléatoire comme modèle sous-jacent aux variations du marché — la distribution des variations dans chacune de ces classes.

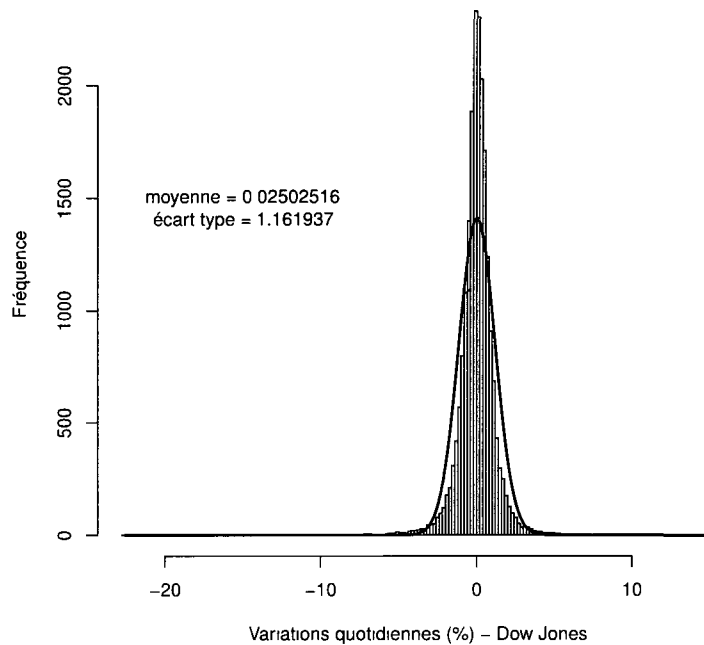


Figure 2.2 : Mouvements du Dow Jones (%) — Octobre 1928 à mai 2010

Première constatation frappante : à environ un demi écart type autour de la moyenne théorique ($x = 0$), les évènements sont beaucoup trop nombreux. Entre un demi ($\frac{1}{2}$) et trois (3) écarts types, les évènements sont légèrement trop peu nombreux. À plus de trois écarts types de la moyenne, les évènements recommencent à être trop nombreux. Il y a une information que le graphique ne confère pas à sa juste valeur : les données aberrantes. L'axe des abscisses semble beaucoup trop large pour le graphique : cela est dû à des évènements comme celui du 19 octobre 1987 où le DJI a encaissé une chute de plus de 22%. Ces données n'apparaissent pas sur le graphique étant donnée l'échelle de l'axe des ordonnées.

Restreignons-nous donc aux variations d'au plus 5%. Nous obtenons les graphiques suivants, selon qu'on choisisse d'utiliser 500 ou 1500 classes pour l'histogramme.

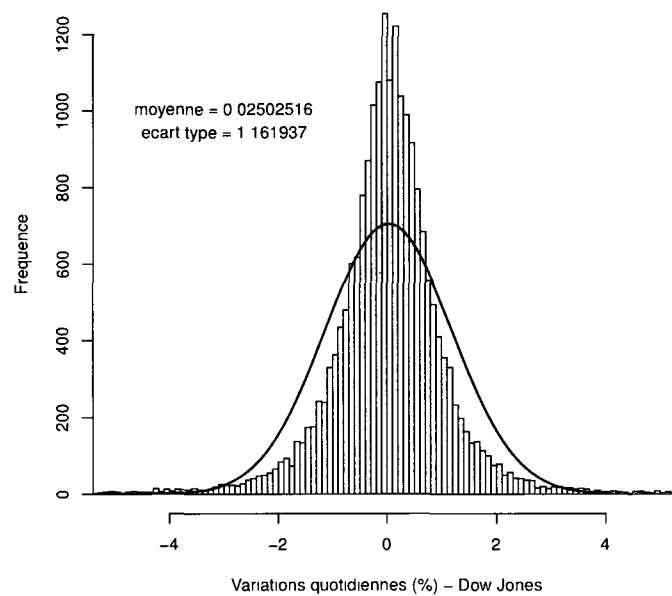


Figure 2.3 : Dow Jones — Octobre 1928 à mai 2010 — 500 classes

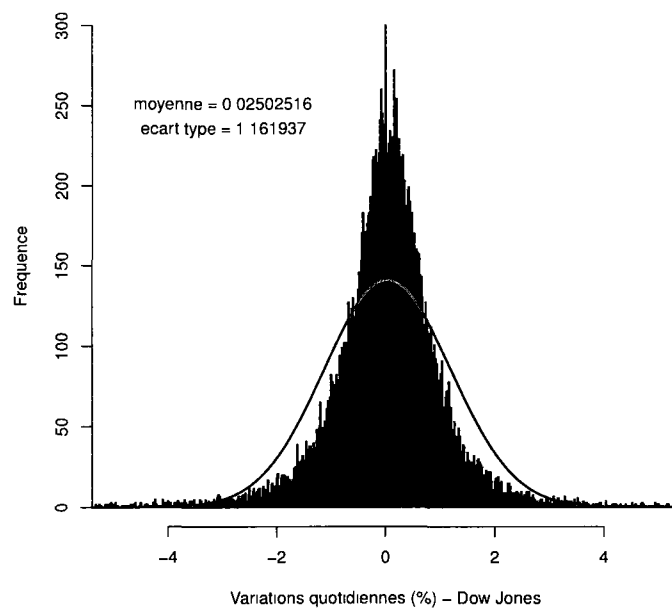


Figure 2.4 : Dow Jones — Octobre 1928 à mai 2010 — 1500 classes

En observant les densités des données⁴ et de la distribution normale théorique correspondante sur un même graphique, nous obtenons :

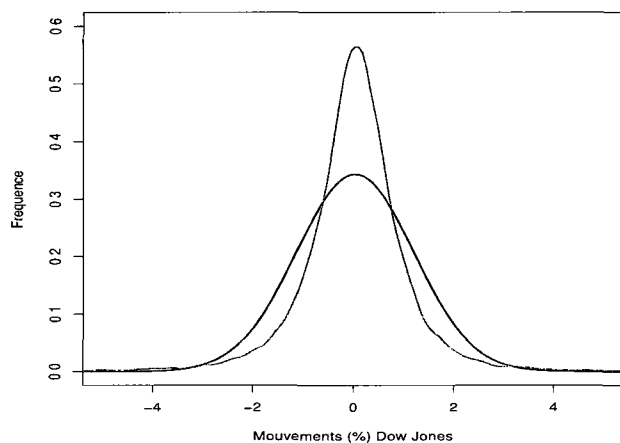


Figure 2.5 : Dow Jones — Octobre 1928 à mai 2010

Il est alors raisonnable de se demander ce qui se passe lorsque nous changeons la fenêtre temporelle. Ainsi, en restreignant la période sous étude de septembre 2008 à mai 2010, nous obtenons le graphique suivant :

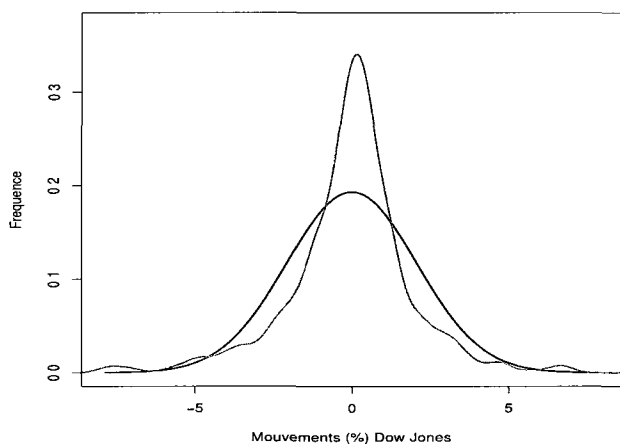


Figure 2.6 : Dow Jones — Septembre 2008 à mai 2010

⁴Il s'agit ici d'une estimation par noyau de la densité.

Mis à part l'échelle, les figures 2.5 et 2.6 sont quasiment identiques. Quel est donc le véritable problème lié à ces distributions ? Il faut réfléchir du point de vue des investisseurs : les petits mouvements ne sont pas susceptibles d'influencer énormément la valeur d'un portefeuille sur une courte période, mais les grandes fluctuations, quant à elles, sont susceptibles de faire exploser la valeur d'un portefeuille, à la hausse comme à la baisse, et ce sur une période de temps relativement restreinte. Ainsi, en présence d'énormes fluctuations, le risque associé à une action quelconque devient lui aussi énorme : c'est ce qui a contribué au succès de la théorie de Markowitz voulant que la diversification d'un portefeuille soit nécessaire pour minimiser le risque associé à ce portefeuille. Ainsi, sur les graphiques précédents, même si les anomalies près de la moyenne et jusqu'à trois écarts types autour de celle-ci sont beaucoup plus notables que l'ampleur des données aberrantes aux extrémités, ce sont ces dernières qui importent vraiment.

La notion de « données aberrantes aux extrémités » de la distribution empirique est ce que les économistes appellent « **queue(s) lourde(s)** »⁵ de la distribution, ou plus communément « **fat tail(s)** » en anglais. Or, la distribution normale — ou « gaussienne », appellation alternative utilisée en l'honneur de Johann Carl Friedrich Gauss — n'est pas une distribution à queue lourde⁶.

La figure suivante illustre cette affirmation en comparant les quantiles des données empiriques à ceux de la distribution normale théorique associée.

⁵La terminologie « queue(s) épaisse(s) » est parfois rencontrée.

⁶La définition mathématique formelle de « distribution à queue(s) lourde(s) » n'étant d'aucune importance dans le présent document.

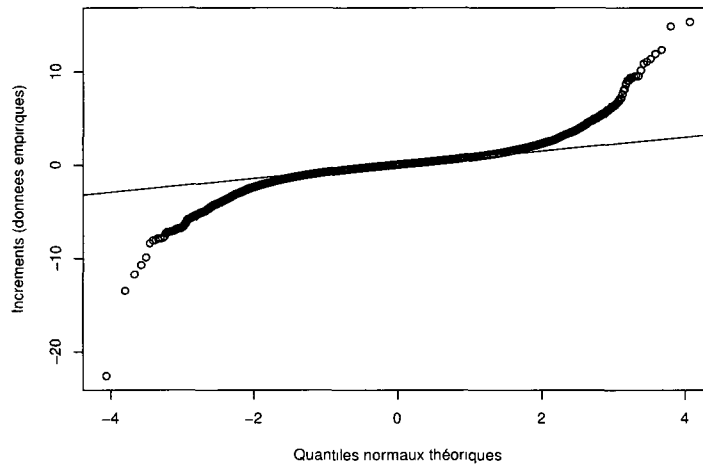


Figure 2.7 : Diagramme de Henry (Dow Jones) — Octobre 1928 à mai 2010

2.2.2 L'indépendance des mouvements et l'amnésie des marchés

Si les fluctuations d'un marché boursier ne peuvent pas être fidèlement décrites par une distribution gaussienne, quelles autres hypothèses du modèle de Bachelier peuvent être remises en question ? Mandelbrot a consacré plusieurs années de sa vie professionnelle à essayer de répondre à cette dernière interrogation.

D'après Mandelbrot, les marchés boursiers ont une mémoire, tant sur des courtes périodes que sur des périodes de plusieurs années. Ses analyses de données financières ont démontré que les variations de prix ont une tendance à se grouper en fonction de leur amplitude. D'une part, lorsque le marché exhibe une tendance à la hausse (respectivement à la baisse), il est plus probable qu'il continue à augmenter (respectivement diminuer). D'autre part, lorsqu'une grosse (respectivement petite) variation survient, d'autres gros (respectivement petits) mouvements sont susceptibles de survenir dans un espace de temps rapproché. Ce phénomène est mieux connu sous le nom de *volatility clustering*⁷ dans la littérature et contredit l'hypothèse

⁷Aucune traduction française de cette expression n'est communément utilisée.

du modèle de Bachelier voulant que les incréments soient indépendants.

Les graphiques suivants montrent l'amplitude des variations quotidiennes du Dow Jones sur des périodes comprises entre 1928 et 2010 ; le comportement quasi-sinusoïdal présent à plusieurs endroits dans les graphiques illustre la tendance des variations à se grouper selon leur amplitude alors que l'asymétrie (par rapport à l'axe $y = 0$) des deux premiers graphiques dénonce la susceptibilité du marché à opter pour une tendance — ou bien croissante, ou bien décroissante — durant cette période.

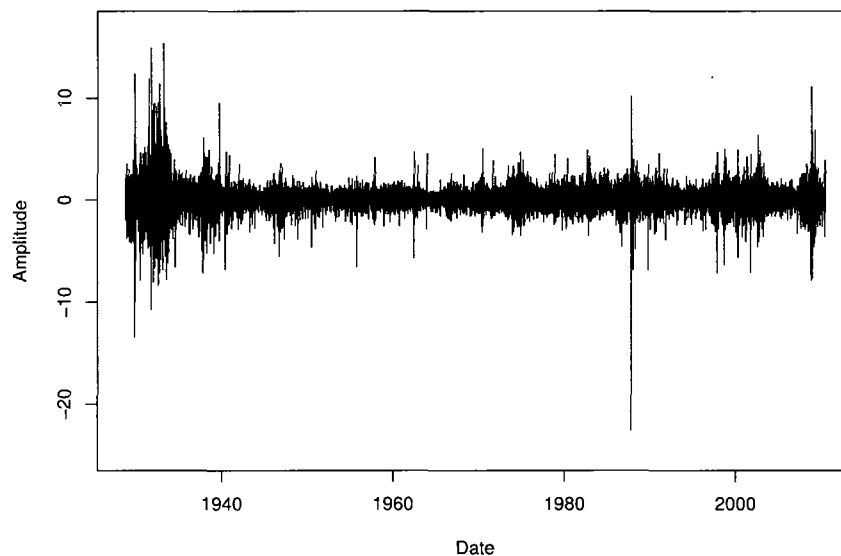


Figure 2.8 : Variations (%) du Dow Jones — 1928 à 2010

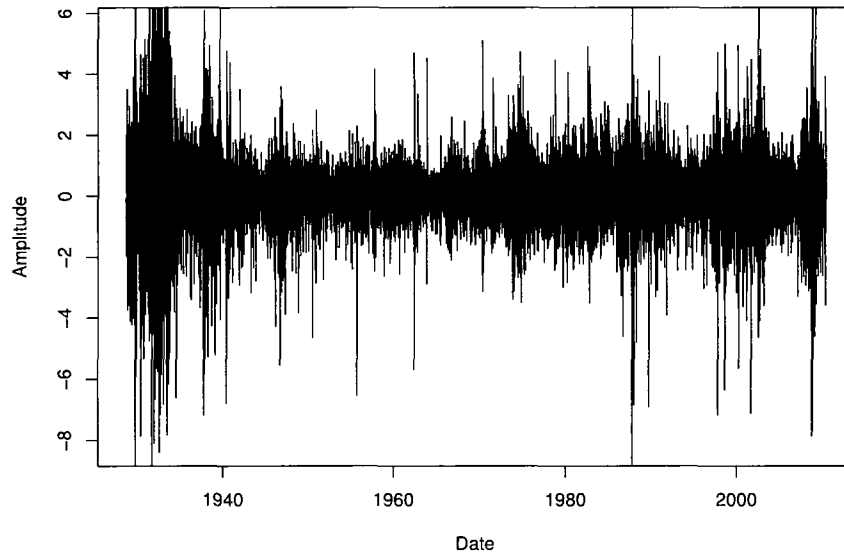


Figure 2.9 : Variations (%) du Dow Jones (zoom) — 1928 à 2010

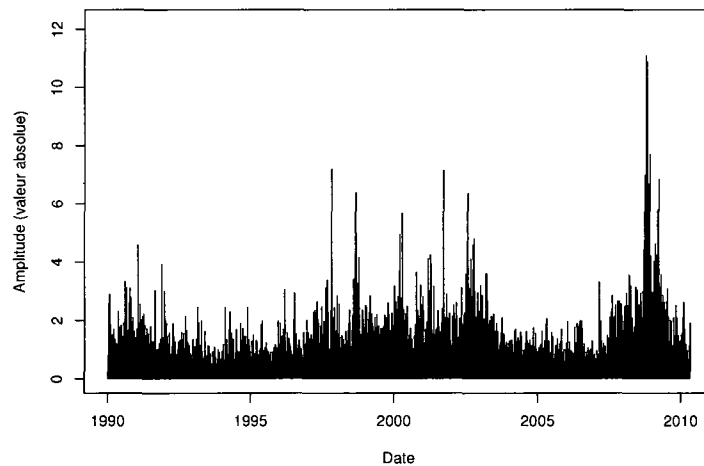


Figure 2.10 : Variations (% , valeur absolue) du Dow Jones — Janvier 1990 à mai 2010

Remarquons sur les graphiques 2.8 et 2.9 que, si toutefois les mouvements d'une certaine amplitude tendent à susciter d'autres fluctuations d'une amplitude similaire sur une période

de temps restreinte, il est également aisé de constater que les signes de ces variations sont souvent opposés, c'est-à-dire que des grosses variations à la hausse tendent à susciter des grosses variations (d'une amplitude similaire) à la baisse dans un avenir relativement rapproché. Mandelbrot dit notamment que *le temps change d'échelle*, un concept connu des négociateurs à la bourse qui utilisent couramment les termes « marché rapide » et « marché lent ». Ainsi, la volatilité est toujours en mouvement dans un marché boursier, mais comment peut-on la modéliser ?

2.3 Les hypothèses du marché

Dans la formulation de modèles économiques, les hypothèses utilisées sont souvent empruntées de modèles précédents et de leurs auteurs. Ainsi, certaines hypothèses deviennent d'usage courant dans la littérature : c'est le cas de *l'hypothèse d'efficience du marché* et des *anticipations rationnelles*.

2.3.1 Hypothèse d'efficience du marché

Exprimée pour la première fois dans la thèse de Bachelier, il fallut attendre la thèse d'Eugene Francis Fama [17][18] publiée en 1965 pour en obtenir une description précise et une théorie complète. L'hypothèse d'efficience du marché (HEM), mieux connue sous son nom anglophone « *Efficient Market Hypothesis* (EMH) », est présente sous trois formes : faible, semi-forte et forte. Toutes les versions stipulent, à leur manière, que les marchés sont efficients d'un point de vue informationnel. Fama et Samuelson sont les principaux instigateurs de cette idéologie.

La version faible de l'HEM stipule que les prix actuels des actions reflètent déjà toute l'information publiquement disponible précédant le moment actuel.

La version semi-forte correspond à la version faible à laquelle on ajoute l'hypothèse que les

prix futurs s'adapteront quasi-instantanément pour refléter les nouvelles informations publiquement disponibles.

La version forte, quant à elle, stipule que les prix des actions reflètent quasi-instantanément toutes les informations, qu'elles soient publiquement disponibles ou non.

En d'autres termes, l'HEM nous dit qu'il est impossible de tirer avantage d'une information (publique) quelconque pour faire un profit sur un marché boursier : en effet, sous cette hypothèse, tous les prix des marchés sont considérés comme étant « justes » et ce à tout moment. De cette hypothèse découle notamment une certaine forme d'indépendance des incréments : puisqu'il est impossible de tirer avantage de toute information jusqu'au moment présent, il en résulte qu'il doit être impossible de prédire les prix futurs étant donnée n'importe laquelle information concernant les prix passés.

Il est intéressant de constater que cette hypothèse était globalement acceptée jusqu'au début des années 1990. Or, dès la fin des années 1960, des études empiriques [48][4][47] soulèvent des doutes quant à la validité de cette hypothèse sur les marchés boursiers.

2.3.2 Anticipations rationnelles

Cette hypothèse, développée par John Fraser Muth [46] en 1961 et popularisée par des économistes prolifiques tel Robert Emerson Lucas Jr, stipule que tous les individus d'un marché effectuent leurs choix en fonction de l'information disponible, de leurs expériences passées et de leur anticipation (rationnelle) du futur, et qu'ils prennent des mesures rationnelles afin que leurs anticipations se réalisent.

Prenons l'exemple d'une compagnie qui, étant donnée l'information publique connue à ce moment, a une raison rationnelle (appelons cette raison R) de croire — ou d'espérer ! — que le prix de son produit X augmentera au cours des prochains mois. Pour en arriver à ses fins, la compagnie décide de diminuer sa production du produit X . De cette manière, la demande reste constante alors que l'offre diminue, causant ainsi une augmentation de la valeur du

produit en question : c'est la mesure rationnelle que la compagnie prend pour veiller à la réalisation de son anticipation. Évidemment, si un évènement imprévu survenait pendant cette période (un scandale, par exemple), la tactique utilisée ne contribuerait en aucune façon à faire augmenter le prix du produit ; l'issue espérée est donc rationnelle uniquement étant donnée la rationalité de R .

En d'autres termes, l'hypothèse des anticipations rationnelles (AR), mieux connue sous son nom anglophone « *Rational Expectations* (RE) », stipule que tous les joueurs actifs d'une économie pensent et agissent de façon rationnelle.

Un cas particulier d'anticipations rationnelles est la théorie des anticipations adaptatives voulant qu'on ajuste nos anticipations en comparant les informations actuelles du marché avec les tendances passées.

Les hypothèses d'anticipations rationnelles et d'efficience du marché sont souvent considérées simultanément : il découle alors des deux hypothèses qu'à tout moment le marché ne possède qu'un seul équilibre prédéterminé et que tous les joueurs actifs forment leurs anticipations rationnellement en fonction de cet unique équilibre et agissent (rationnellement) en conséquence — remarquons ici le lien avec l'idée de Sharpe. Or, l'hypothèse d'AR est contestée par bon nombre de personnes étudiant les implications sociales et psychologiques de la dynamique interne des marchés : en particulier, plusieurs facteurs font en sorte que les décisions prises par les investisseurs ne sont pas rationnelles. L'être humain, doit-on le rappeler, n'est lui-même pas rationnel ; peut-on vraiment s'attendre à ce que ses décisions le soient, surtout lorsqu'il est question de richesse ou de ruine ?

2.3.3 Autres modèles

Évidemment, le but de ce chapitre n'était pas d'établir une liste exhaustive de tous les modèles proposés au cours du dernier siècle, mais bien de fournir une idée générale de l'évolution des divers modèles utilisés durant cette période et d'exposer certains problèmes liés à leur utilisation.

tion. En effet, plusieurs autres modèles (ou théories) ont été proposés dont quelques-uns sont encore couramment utilisés. L'arrivée du CAPM aura d'ailleurs donné lieu à l'élaboration de deux autres modèles très importants qui n'ont pas été abordés dans cet ouvrage : le modèle de Fama-French et le *Arbitrage Pricing Theory (APT)*⁸.

⁸La traduction française étant « modèle d'évaluation par arbitrage (MEA) ».

CHAPITRE 3

Modèles contemporains

Il est passablement évident à la vue des graphiques de la section précédente qu'une distribution normale est inappropriée pour décrire les variations de marchés boursiers¹. Mais alors quelle distribution suivent-elles ? Comment peut-on les modéliser ?

Nous discuterons dans ce chapitre des problèmes inhérents à l'étude des marchés financiers et de la disparité des opinions régnant encore aujourd'hui dans le monde de la finance.

3.1 Problèmes inhérents à l'étude des marchés

Il existe quelques problèmes concernant l'analyse et la modélisation d'un marché auxquels il n'existe pas de réponse unique et absolue. Même en considérant les concepts les plus fondamentaux, des problématiques d'envergure peuvent être soulevées.

3.1.1 Qu'est-ce que le risque ?

Pourquoi investir à la bourse ? Pourquoi ne pas tout simplement se limiter aux obligations d'État, qui elles fournissent un taux d'intérêt fixe ? Évidemment, c'est parce qu'un inves-

¹Les graphiques d'autres marchés boursiers présentent les mêmes caractéristiques que ceux du Dow Jones.

tisseur a la possibilité de réaliser un plus gros profit à la bourse. Mais si c'est le cas, c'est uniquement dû au risque associé : le marché est risqué et par conséquent il doit offrir la possibilité de faire plus d'argent que les obligations d'État — qui elles ne sont pas du tout risquées —, sans quoi il n'y a absolument aucun attrait ! En particulier, plus le risque associé à l'investissement est élevé, plus le taux de retour potentiel associé à cet investissement doit être élevé. Ainsi, la valeur d'une action doit être déterminée en fonction du risque qui y est associé. Et alors, comment mesure-t-on ce risque ?

Étant donnée une action, le risque associé doit représenter son potentiel de rendement. Or, la valeur d'une action ne peut évoluer qu'au rythme des fluctuations du marché. Ainsi, d'entre toutes les actions, celle à laquelle sera adjoint le plus grand risque est celle dont la valeur réagira le plus violemment à une même fluctuation du marché : c'est ce que nous appelons la *volatilité* de l'action [1] (ou du processus stochastique sous-jacent). C'est notamment l'idée derrière le β de Sharpe.

Toutefois, la question de mesurer ladite volatilité d'une action n'en est pas une facile : en effet, la volatilité d'une action dépend du modèle que l'on choisit d'utiliser. Supposons que l'on choisisse de modéliser le prix d'une action à l'aide d'un processus de Wiener. Ainsi, les incréments sont normalement distribués et donc la volatilité correspond à l'écart type (ou la variance) — qui varie en fonction du temps d'un point de vue empirique, mais qui devrait se stabiliser vers la valeur théorique sur une longue période — de cette distribution normale. Mais qu'en est-il du risque ? Supposons que l'on considère une option européenne — soit une option selon laquelle un individu a la possibilité d'acheter un nombre de parts d'une action à un temps futur T prédéterminé — sur cette même action. Bien qu'en espérance on puisse s'attendre à ce que le prix de l'action soit demeuré inchangé au temps T (puisque'il s'agit d'un mouvement brownien, qui est une martingale), il va de soi que plus ce temps T est loin dans le futur, plus la valeur de l'action a la possibilité de s'écarter de la valeur actuelle : le risque associé à cette option dépend donc également de la période de temps couverte par l'option.

Néanmoins, ce sera le modèle pour lequel nous opterons qui dictera le risque associé à un

investissement : c'est la raison primordiale pour laquelle nous cherchons à modéliser les marchés. La méthode standard actuelle consiste à observer les prix des options fixés par les forces du marché et de calculer la volatilité implicite en inversant la formule de Black-Scholes.

3.1.2 Comportement interne ou dynamique globale ?

Afin d'essayer de développer un modèle adéquat pour modéliser un marché financier, il faut s'arrêter quelques minutes et réfléchir à son fonctionnement. Il s'agit d'un regroupement d'individus et de sociétés qui spéculent quant aux prix et aux tendances futurs du marché, et qui prennent des décisions rationnelles d'investissement en fonction de leurs pressentiments, de leurs expériences passées et de l'information dont ils disposent. Mais comment cela fonctionne-t-il ? Comment une information nouvelle circule-t-elle auprès des investisseurs et à quelle vitesse ? En présence d'une nouvelle information, un individu spécifique décide-t-il de changer de tactique ou non, et pourquoi ? Est-ce que tous les investisseurs affichent le même profil d'investissement ? Deux parties présentant le même profil prennent-elles les mêmes décisions étant données les mêmes informations ? Les décisions de l'une affectent-elles les décisions d'une autre et, le cas échéant, à quel rythme ?

Au cours des années 1990 naquit le domaine de l'*éconophysique*, une branche étudiant l'application de modèles, techniques et théories développées dans le monde de la physique à des problèmes concrets de l'économie. Avec l'apparition d'un tel domaine, la question de modéliser un marché financier a pris un tout autre sens : alors qu'on s'intéressait jusqu'alors principalement à la dynamique régissant l'évolution du marché en tant que système global, l'arrivée de l'éconophysique a poussé des mathématiciens à vouloir modéliser le fonctionnement interne des composantes du système en dictant les règles qui régissent leur comportement. Bien que ceci soit une approche viable en physique, étant donné le caractère quasi-homogène des constituants d'un système et le nombre généralement limité de règles régissant leurs interactions, elle est beaucoup plus difficilement applicable au monde de la

finance. En effet, l'interaction entre les diverses composantes d'un marché financier relève plus souvent qu'autrement de la psychologie et les règles dictant ces comportements sont si complexes qu'il nous est impossible de les modéliser mathématiquement, sans compter bien sûr l'interdépendance qui peut exister entre ces différentes règles.

Il est donc intéressant de se questionner : si le comportement interne du système est si complexe, nous est-il possible de modéliser la dynamique globale de ce système en ignorant complètement l'interaction de ses divers constituants ?

3.1.3 Autorégulation du marché

Une particularité intéressante d'un marché financier est qu'il existe un processus de rétroaction (ou *feedback*), tant positive que négative. L'état actuel du marché pousse les individus à réagir de façon à ramener le marché à un état d'équilibre. À titre d'exemple, si le prix d'une action est trop bas, les individus remarqueront le déséquilibre et agiront en conséquence en achetant des parts de cette action ; ceci aura pour effet d'augmenter le prix de l'action. En d'autres termes, les règles dictant la dynamique interne du système ont pour effet de rapporter le marché à un état d'équilibre. C'est la *rétroaction négative* du système, car elle a pour effet d'amoindrir le déséquilibre.

Cependant, lorsque la dynamique du système global change, les règles dictant la dynamique interne changent également : lorsqu'une bulle économique apparaît, il y a spéculation. Une étude de Andreassen et Kraus (1988) [2] a cherché à explorer les tendances d'investissement des individus. À l'aide de données financières réelles, Andreassen et Kraus ont demandé à des gens ayant une connaissance de l'économie d'investir (de façon fictive) dans le marché. Initialement, les individus sont mis au courant de la situation actuelle du marché et les données financières leur sont fournies une par une : à chaque nouvelle donnée, on demande aux investisseurs de rectifier leur tactique d'investissement (au besoin). Évidemment, les décisions des individus n'affectent pas les données financières qui leur sont fournies, mais étant réelles, ces données incorporent déjà un processus de rétroaction. Les résultats sont frappants : lorsque les

variations de prix sont relativement constantes (c'est-à-dire que la variance est relativement fixe) sur une période donnée, les investisseurs achètent lorsque les prix baissent et vendent lorsque les prix augmentent. Toutefois, lorsque la dynamique du marché change et que, sur une période donnée, les prix exhibent des tendances (ou bien à la hausse, ou bien à la baisse), les investisseurs inversent leur tactique : ils achètent lorsque les prix augmentent et vendent lorsque les prix diminuent.

Lorsque le marché exhibe une tendance, il y a souvent spéculation : les investisseurs rationnels, sous la nouvelle que les prix augmenteront demain, en profitent pour acheter beaucoup de parts aujourd'hui, ce qui a pour effet de faire augmenter le prix de l'action plus haut que prévu. En réaction à cette hausse de prix imprévue, les spéculateurs continuent à acheter davantage de parts dans l'espoir que la tendance à la hausse se maintienne, alors que certains investisseurs (qui ont généralement contribué à alimenter la spéculation initiale) en profitent pour vendre les leurs, question de faire un profit. Néanmoins, la tendance apparaît et continue jusqu'à ce que les investisseurs rectifient la situation ou jusqu'à ce qu'il y ait un krach : c'est ce que nous appelons une *bulle spéculative* ou *bulle économique*. Ce processus est la *rétroaction positive* du système, car elle a pour effet d'augmenter le déséquilibre.

Qu'importe le type de rétroaction, il y a néanmoins un problème : les spéculations et les actions des individus formant le système ont un impact direct sur le système. L'issue du système, à son tour, a un impact sur la dynamique interne du système (règles de décision des investisseurs). Comment donc doit-on modéliser le système pour incorporer un processus de rétroaction ? Ce processus peut-il être modélisé au sein d'un système d'équations différentielles décrivant le système global ou alors doit-on avoir recours à un processus itératif ?

3.1.4 L'unicité des données

Il n'est pas surprenant que le domaine de l'économie emprunte des outils au monde de la physique ; un marché boursier est un système dynamique chaotique s'apparentant à plusieurs phénomènes physiques traitant de l'interaction entre les particules et à d'autres phénomènes

naturels dont la volcanologie et l'hydrologie. Toutefois, il y a une particularité délicate des marchés qui n'est pas partagée avec les phénomènes physiques : l'impossibilité de reproduire l'expérience. S'il nous est possible de reproduire une expérience physique autant de fois que le temps et l'argent nous le permettent, nous n'avons qu'une et une seule réalisation d'un marché boursier. Cette particularité rend l'analyse d'un marché spécifique plutôt délicate et sa modélisation encore plus difficile : en effet, une seule réalisation d'un processus stochastique est nettement insuffisante pour espérer y déceler des tendances et l'impossibilité de comparer nos observations avec une autre réalisation rend quasi impossible la tâche de distinguer les événements et les tendances dus au hasard de ceux typiques du modèle sous-jacent. En particulier, la volatilité élevée du marché ces dernières années est-elle due à une tendance aléatoire qui s'est présentée il y a déjà un certain temps ou est-ce que toutes les réalisations du marché (s'il était possible d'en avoir plusieurs) présenteraient cette même particularité temporelle, nonobstant leurs différences jusqu'alors ? Ou encore est-elle due à une tendance passée qui n'était pas le résultat d'un hasard, mais bien de celui du véritable tempérament du marché ?

Bien sûr, il y a plusieurs marchés financiers à travers le monde, ce qui pourrait laisser croire que nous sommes en présence d'un nombre adéquat d'observations. Or, rien ne nous garantit que ces marchés soient modélisables d'une manière commune et universelle, et l'interdépendance de tous ces marchés nous donne l'assurance que les observations provenant de divers marchés ne seraient pas indépendantes : les fluctuations de l'un sont souvent représentatifs des fluctuations des autres.

3.1.5 Qu'est-ce qu'une tendance

Il est aisé pour un être humain de regarder le graphique des variations d'un marché et d'y déceler une « tendance », à la hausse comme à la baisse, sur une période de temps donnée. Mais comment peut-on définir ce concept mathématiquement ? Même si cette question semble simple et anodine, elle recèle plusieurs subtilités et ses implications sont énormes

dans l'analyse du marché : plusieurs stratégies employées par les investisseurs incorporent la notion de « tendance ».

Restreignons-nous à l'étude d'une période fixée. Comment déterminer, sur cette période, si le marché affiche une tendance haussière ou baissière. L'idée la plus simple serait de comparer le nombre de mouvements vers le haut au nombre de mouvements vers le bas. Or cette approche présente déjà une difficulté : elle ne tient pas compte des intensités des mouvements. En effet, il se pourrait que deux variations sur trois présentent une augmentation d'une unité alors que chaque troisième variation serait une diminution de 5 unités. Ainsi, il y aurait près de deux fois plus de mouvements ascendants, mais la tendance générale serait à la baisse.

Considérons alors l'intensité des mouvements : intéressons-nous à la somme de toutes les fluctuations. Ainsi, si le total est positif, il y aura eu davantage de variations à la hausse que de variations à la baisse ; et inversement si le résultat est négatif. Une nouvelle question est soulevée : à partir de quelle valeur peut-on conclure qu'il s'agit bel et bien d'une tendance ? Cette valeur doit-elle être constante ou devrait-elle être une fonction de la largeur de la fenêtre temporelle utilisée ? Doit-on considérer la volatilité ? Visiblement, il ne semble pas y avoir de réponse simple.

De nos jours, certains économistes affirment encore qu'il est impossible de modéliser et de prévoir des données financières issues de distributions non-stationnaires. Or, il s'agit, à mon avis, d'un paralogisme issu d'une mauvaise perception de la non-stationnarité : bien qu'il soit beaucoup plus difficile de trouver un modèle entièrement adéquat pour modéliser une telle distribution et qu'il soit d'autant plus difficile d'en vérifier la validité, la non-stationnarité peut à certains moments nous offrir de l'information additionnelle par l'entremise des tendances qui en sont issues. C'est d'ailleurs l'idée dont tire avantage le modèle LPPL abordé dans le chapitre 4 du présent document. Bien sûr, il reste toujours le problème de définir le type de stationnarité dont il est question.

3.1.6 Qu'est-ce qu'une bulle économique

Lorsque nous cherchons à développer un modèle pour les krachs boursiers issus de bulles économiques, il faut d'abord se demander « qu'est-ce qu'une bulle » ? Encore une fois, la question est bien simple alors que la réponse n'est pas du tout claire mathématiquement.

Nous pourrions décider de définir une bulle économique comme étant une période affichant une tendance à la hausse suivie d'une chute de prix très rapide. Mais alors quelles doivent être les amplitudes de la hausse et de la chute afin que nous considérions le tout comme étant une bulle ? La portion « chute » a-t-elle alors vraiment de l'importance ? Peut-il y avoir une bulle sans krach ? Combien rapide doit être la chute afin qu'il s'agisse bel et bien d'un krach et non pas tout simplement d'une tendance à la hausse suivie d'une tendance à la baisse ? De plus, cette caractérisation d'une bulle emprunte la notion de « tendance » qui, rappelons-le, est en soi une notion difficile à décrire mathématiquement.

Ainsi, nous cherchons à modéliser une bulle économique sans être capable de la définir adéquatement. En effet, différents chercheurs ne s'entendent pas sur la définition mathématique d'une bulle : Brée et Joseph, qui ont effectué leur propre analyse du modèle LPPL proposé par Sornette *et al.*, réussissent à détecter davantage de bulles économiques sur une période fixée que ne le fait l'étude originale. Bien que les deux études semblent dénombrer les bulles en détectant des hausses et des baisses de prix d'une certaine amplitude sur un intervalle de temps fixé, elles ne s'entendent guère sur l'intervalle à utiliser et l'amplitude à considérer.

3.2 À la recherche d'une distribution appropriée

Avec les différentes idéologies gouvernant le monde de l'économie, nous avons vu naître, au cours des dernières décennies, diverses nouvelles branches : l'économétrie, l'éconophy-sique, la finance comportementale, etc. Différentes façons de penser qui tentent d'expliquer ou d'élucider, chacun à leur manière, le fonctionnement ou la dynamique d'un marché. Alors

que l'un propose de dicter l'ensemble des règles régissant le fonctionnement interne du système, un autre prône de déjouer cette problématique en énonçant uniquement les équations relatives à la dynamique globale. Et si l'un tente de justifier le fonctionnement interne du système en le comparant à des phénomènes physiques ou naturels, un autre cherche à l'expliquer par l'entremise de la psychologie behaviorale.

Mais, d'entre toutes les manières de procéder qui nous sont proposées, en existe-t-il réellement une qui soit *meilleure* que les autres ?

3.2.1 Continuité du marché

Étant issus d'un mouvement brownien, les incréments des marchés boursiers devraient suivre une distribution continue. Or, dans la vie réelle, l'ensemble des transactions effectuées est mesurable sur un domaine discret : il est donc naturel de se demander si la modélisation d'un marché boursier par un processus continu est raisonnable.

Mandelbrot argumente que deux graphiques illustrant les variations d'un marché boursier sur deux périodes différentes — par exemple, l'un couvrant une période d'une semaine et l'autre présentant les fluctuations sur vingt ans — sont indiscernables l'un de l'autre. Le nom qu'il a choisi d'attribuer à cette propriété est « l'autosimilarité »² : cette notion est à la base des mathématiques fractales, théorie qu'il publiera vers le milieu des années 1970 suite à l'élaboration de ses modèles financiers.

²« *Self-similarity* » en anglais.

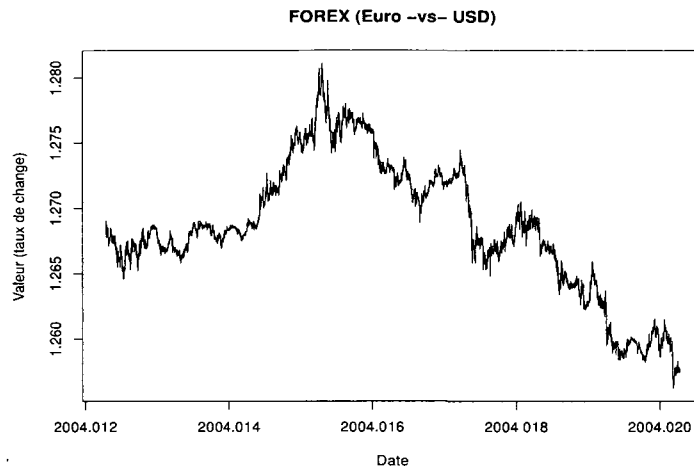


Figure 3.1 : FOREX — Une journée

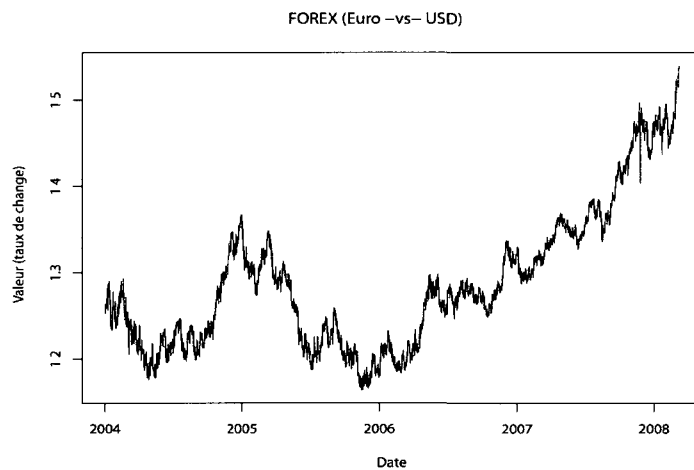


Figure 3.2 : FOREX — Janvier 2004 à mars 2008

Or, il existe des limites à l'autosimilarité : à très petite échelle, il se produit un effet de discrétisation. Il s'agit d'une forme d'invariance d'échelle définie uniquement sur un domaine borné. En présence de données financières, lorsque nous abordons une période de cinq minutes, par exemple, il nous est possible de voir absolument tous les sommets véritables du graphique : ce dernier ressemble davantage à une marche aléatoire renormalisée qu'à un mouvement brownien. En contrepartie, lorsque la période sous étude est très grande (200 ans, par

exemple), l'image n'est pas définie sur le domaine complet, mais le graphique de la portion définie ressemble davantage à un mouvement brownien qu'à une marche aléatoire renormalisée : relativement à l'échelle de la période sous étude, la distribution semble continue. Ces diverses constatations quant à la discrétisation (ou non) du domaine temporel ainsi que l'auto-similarité (bien que limitée) des graphiques ont poussé Mandelbrot à reconsidérer ses propres hypothèses et à formuler plusieurs modèles à travers les années.

Outre Mandelbrot, les mathématiciens ne s'entendent pas tous sur la question : alors que les défenseurs de la théorie classique maintiennent que la distribution des observations d'un marché boursier est continue — ou plutôt que les données réelles observées sont le résultat d'observations d'un modèle continu sous-jacent effectuées sur un domaine discret —, les partisans du renouveau théorique en économie sont plutôt d'avis qu'il existe des variations d'une amplitude si grande relativement à la volatilité avoisinante qu'il est inconcevable que ces données puissent provenir d'un processus purement continu. En particulier, le 6 mai 2010, les marchés boursiers à travers le monde ont subi une chute quasi-instantanée d'une envergure extrêmement improbable : alors que le Dow Jones a présenté une baisse d'environ 7% sur une période de 15 minutes, le Standard & Poor's 500 (S&P500) et le Nasdaq Composite ont encaissé des chutes d'un peu plus de 3% sur une période similaire. Laquelle des deux hypothèses suivantes semble la plus probable : s'agit-il d'un saut dans la distribution des variations de chacun de ces marchés (cas discret) ou alors d'un changement radical dans la volatilité de ces distributions (cas continu) ?

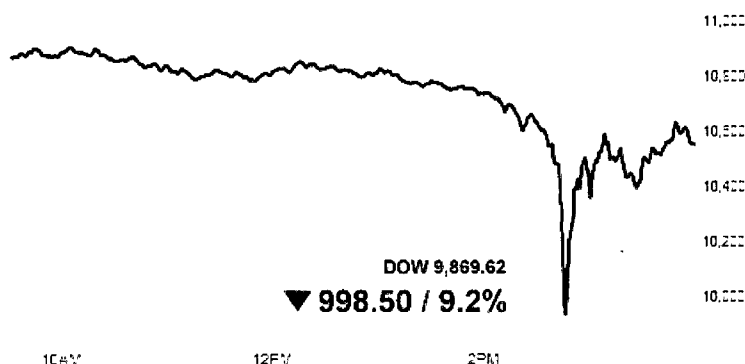


Figure 3.3 : Chute du Dow Jones — 6 mai 2010

3.2.2 Engle et ARCH

Suite aux observations voulant que les variations d'un marché aient une tendance à se grouper en fonction de leur amplitude, Robert Fry Engle III propose en 1982 [16] une façon de modéliser la volatilité d'une action à travers le temps. Ce modèle, mieux connu sous l'acronyme « **ARCH** » (*Autoregressive Conditional Heteroskedasticity*)³, est une forme particulière d'un processus autorégressif (discret), c'est-à-dire une série temporelle dont la valeur actuelle est expliquée par les valeurs passées du processus (et ce à tout moment). Engle a eut l'idée de modéliser la volatilité actuelle d'une action en fonction des volatilités passées. Plutôt que d'utiliser les volatilités directement, Engle propose d'estimer les erreurs ϵ_t des observations par rapport à un modèle de comparaison théorique et de modéliser l'erreur actuelle en fonction des erreurs passées. Décrivons le modèle ARCH.

Définition 3.2.1

Soit $\{\eta_t\}_{t \geq 1}$, une série temporelle. L'opérateur retard L associé à la série $\{\eta_t\}$ est l'opérateur qui associe à tout élément η_t l'observation précédente η_{t-1} , c'est-à-dire $L\eta_t = \eta_{t-1}$ pour tout $t > 1$.

Définition 3.2.2

Soit L l'opérateur retard associé à une série temporelle $\{\eta_t\}_{t \geq 1}$. Un polynôme retard $\alpha(L)$ d'ordre p est un polynôme en L , c'est-à-dire

$$\alpha(L) = \sum_{i=1}^p \alpha_i L^i$$

où L^i dénote la composition de l'opérateur L avec lui-même i fois.

Exemple 3.2.3

Selon ces notations, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_p)$ représente un vecteur à p éléments. Ainsi, pour

³Une traduction possible est « hétéroscédasticité conditionnelle autorégressive ».

$p = 4$ et une série temporelle $\{\eta_t\}_{t \geq 1}$, nous aurions :

$$\begin{aligned}
\alpha(L)\eta_t &= \left(\sum_{i=1}^4 \alpha_i L^i\right)\eta_t \\
&= (\alpha_1 L + \alpha_2 L^2 + \alpha_3 L^3 + \alpha_4 L^4)\eta_t \\
&= \alpha_1 L\eta_t + \alpha_2 L^2\eta_t + \alpha_3 L^3\eta_t + \alpha_4 L^4\eta_t \\
&= \alpha_1 L\eta_t + \alpha_2 (L \circ L)\eta_t + \alpha_3 (L \circ L \circ L)\eta_t + \alpha_4 (L \circ L \circ L \circ L)\eta_t \\
&= \alpha_1 \eta_{t-1} + \alpha_2 L\eta_{t-1} + \alpha_3 (L \circ L)\eta_{t-1} + \alpha_4 (L \circ L \circ L)\eta_{t-1} \\
&= \alpha_1 \eta_{t-1} + \alpha_2 \eta_{t-2} + \alpha_3 L\eta_{t-2} + \alpha_4 (L \circ L)\eta_{t-2} \\
&= \alpha_1 \eta_{t-1} + \alpha_2 \eta_{t-2} + \alpha_3 \eta_{t-3} + \alpha_4 L\eta_{t-3} \\
&= \alpha_1 \eta_{t-1} + \alpha_2 \eta_{t-2} + \alpha_3 \eta_{t-3} + \alpha_4 \eta_{t-4}
\end{aligned}$$

pour $t > 4$. En général, si $p = k$, nous avons donc

$$\alpha(L)\eta_t = \alpha_1 \eta_{t-1} + \alpha_2 \eta_{t-2} + \dots + \alpha_p \eta_{t-k}$$

pour $t > k$.

Ainsi, dans ses travaux de 1982, Engle propose d'utiliser les modèles de la classe ARCH définis comme suit.

Définition 3.2.4

Soit $\{\eta_t\}_{t \geq 1}$ une série temporelle telle que $\eta_{t+1} \sim \mathcal{N}(0, \sigma_t^2)$ conditionnellement à l'information au temps t pour tout $t \geq 1$ et soit $\alpha(L)$ un polynôme retard de degré p à coefficients non-négatifs associé à $\{\eta_t\}$. Un modèle de type ARCH(p) est un modèle de la forme

$$\sigma_t^2 = \omega + \alpha(L)\eta_t^2, \quad (3.1)$$

où $\omega \geq 0$ et σ_t^2 est la variance conditionnelle⁴ de η_{t+1} étant donnée l'information au temps t .

Exemple 3.2.5

Un modèle de type ARCH(1) est de la forme $\sigma_t^2 = \omega + \alpha_1 \eta_{t-1}^2$, où $\omega, \alpha_1 \geq 0$.

⁴De façon équivalente, σ_t^2 est l'espérance de η_{t+1}^2 conditionnelle à l'information au temps t .

Ainsi, en supposant les incréments d'un marché boursier comme étant normalement distribués de moyenne zéro et de variance σ_t^2 au temps t , il est possible d'appliquer le modèle ARCH. En particulier, Campbell, Lo, MacKinlay et Lo [11] démontrent qu'une telle modélisation fait apparaître des queues plus lourdes qu'une distribution normale lorsque les variances σ_t varient dans le temps. Le modèle accentue également la distribution des incréments autour de la moyenne. Or, bien qu'il s'agisse d'une amélioration, plusieurs études ont démontré que les queues issues d'une modélisation à l'aide de ARCH sont moins lourdes que les données empiriques recueillies à travers les années sur divers marchés boursiers.

En 1986, Tim Peter Bollerslev propose une généralisation du modèle d'Engle [6] permettant d'éviter l'estimation d'un grand nombre de coefficients dans les polynômes retard d'ordres supérieurs. Ce modèle, mieux connu sous l'acronyme « **GARCH** » (*Generalized Autoregressive Conditional Heteroskedasticity*)⁵, utilise la formule suivante :

$$\sigma_t^2 = \omega + \beta(L)\sigma_{t-1}^2 + \alpha(L)\eta_t^2, \quad (3.2)$$

où $\omega \geq 0$ et $\beta(L), \alpha(L)$ sont des polynômes retard à coefficients non-négatifs de degrés p et q (respectivement) associés aux séries $\{\sigma_t^2\}$ et $\{\eta_t^2\}$ (respectivement). Un tel modèle est noté par $GARCH(p, q)$.

Plusieurs autres variations du modèle existent, notamment NGARCH, IGARCH, EGARCH, QGARCH, FGARCH, FIGARCH, FIEGARCH, TGARCH, GARCH-M, GJR-GARCH, etc. [21]

Parmi les plus grandes difficultés rattachées à l'utilisation de ces méthodes se trouvent la détermination du nombre de termes (p et q) à utiliser ainsi que l'estimation des coefficients des polynômes retard.

⁵Une traduction possible est « hétéroscédasticité conditionnelle autorégressive généralisée ».

3.2.3 Mandelbrot et ses multiples modèles

Alors qu'il écrit sa thèse de doctorat, Mandelbrot consulte les ouvrages de George Kingsley Zipf et Vilfredo Pareto. Ces ouvrages suggèrent que les distributions associées à plusieurs phénomènes naturels et scientifiques suivent une **loi de puissance**. Une telle loi est donnée par :

$$f(x) = cx^\alpha + \varepsilon, \quad (3.3)$$

où c et α sont des constantes, et ε est une fonction asymptotiquement petite⁶ de l'ordre maximal $o(x^\alpha)$. De façon plus courante, ε est une constante.

Restreignons-nous au cas $\varepsilon = 0$, quitte à translater toutes les données en ordonnée. Nous avons donc la forme restreinte $f(x) = cx^\alpha$. Ainsi, en prenant le logarithme, nous obtenons :

$$\log_\alpha f(x) = \alpha \log_\alpha(x) + \log_\alpha(c), \quad (3.4)$$

d'où $\log_\alpha f(x)$ est une droite de pente α . Afin de vérifier si une distribution suit une loi de puissance, il suffit donc de la tracer à l'aide d'une échelle logarithmique et de vérifier si le graphe résultant est une droite.

Dans ses ouvrages, Pareto affirme que la queue de la dispersion de la richesse dans une société correspondant aux revenus élevés suit une telle loi. Or, il s'agit en fait d'une loi de puissance inverse, c'est-à-dire une loi de puissance dont l'exposant est strictement négatif. Ainsi, si X est une variable aléatoire représentant le revenu d'un individu dans la société, nous avons

$$\mathbb{P}(X > x) \sim x^{-\alpha}. \quad (3.5)$$

Plus explicitement, cette dernière équation stipule que $\mathbb{P}(X > x) = cx^{-\alpha} + \varepsilon$ où c et α sont des constantes, et ε est un terme d'incertitude. Ainsi, étudiant les prix du coton dans un laboratoire d'IBM, Mandelbrot adopte cette idée et tente de l'appliquer à ses propres données.

⁶C'est-à-dire $\lim_{x \rightarrow \infty} \frac{\varepsilon}{x^\alpha} = 0$.

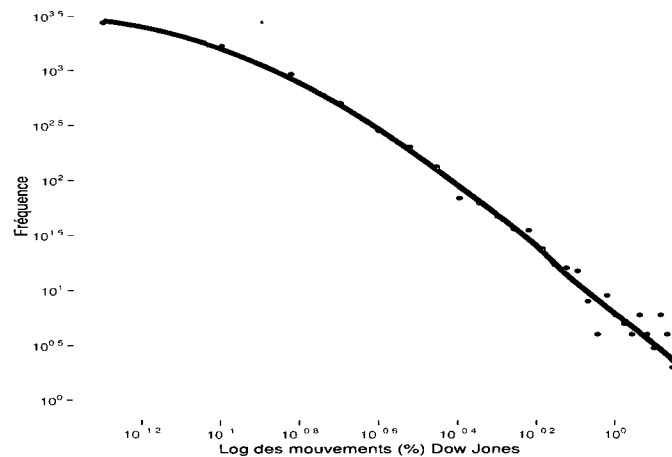


Figure 3.4 : Dow Jones — Logarithme des classes d'incrément positifs

Sur le graphique précédent, nous constatons en effet que, bien qu'il s'agisse d'une courbe près de la moyenne de la distribution (c'est-à-dire près de zéro), une droite apparaît clairement en s'approchant de la queue. Suite à cette constatation, Mandelbrot décide d'étudier l'application des lois de puissance inverses aux marchés boursiers.

Au cours des années, Mandelbrot consacre une énorme partie de son temps à analyser des données financières de divers marchés et à formuler des hypothèses. En fonction des conclusions de ses études (et de celles d'autres chercheurs), il développe divers modèles à travers les années qu'il adaptera graduellement en fonction des résultats que lui et le reste du monde scientifique établiront. C'est ainsi que nous voyons naître en 1963 le modèle mesofractal [34], en 1965 le modèle unifractal [35] et en 1972 le modèle multifractal [36]. Ce dernier modèle sera rectifié en 1997 par Mandelbrot lui-même [41][39].

Pour Mandelbrot, un marché est une entité à part entière qui possède un comportement et qui exhibe certaines caractéristiques. Il consacre donc plusieurs années à étudier les particularités des marchés, à formuler des hypothèses et à mettre ces dernières à l'épreuve. Les deux particularités qu'il semble chérir le plus sont *l'invariance d'échelle* et *l'indépendance des incréments*.

Au cours des premières années, Mandelbrot constate que la distribution des incréments semble coïncée entre une distribution normale et une distribution de Cauchy. Il remarque que ces deux distributions sont des cas particuliers d'une classe beaucoup plus large : les **distributions L-stables**⁷. Le modèle mesofractal proposé par Mandelbrot en 1963 est d'ailleurs une conséquence du fait de supposer que les incréments sont indépendants et distribués selon une loi L-stable plutôt que gaussienne.

Or l'hypothèse d'indépendance ne plaît pas particulièrement à Mandelbrot, qui présente des arguments et des travaux montrant qu'il existe une certaine forme de mémoire au sein d'un marché boursier. En particulier, il insiste sur *l'autosimilarité* et *l'autoaffinité* des graphiques [40]. Alors que l'autosimilarité suppose une similitude issue d'un changement d'échelle isotrope (identique sur les deux axes), l'autoaffinité suppose une similitude issue d'un changement d'échelle anisotrope (différent sur les deux axes) [37].

Définition 3.2.6

Un processus stochastique $\{X_t\}_{t \geq 0}$ est dit **autoaffine** s'il existe un nombre réel H tel que

$$X(ct) = c^H X(t) \quad (3.6)$$

pour tout nombres réels c et $t \geq 0$. Le nombre H est appelé **l'index d'autoaffinité**.

Tous les modèles proposés par Mandelbrot se basent sur l'autoaffinité et le plus récent modèle englobe les précédents sous forme de cas particuliers. Afin d'incorporer la notion de mémoire, Mandelbrot et van Ness [43] introduisent la notion de *mouvement brownien fractionnaire*.

Définition 3.2.7

Un processus stochastique $\{X_t\}_{t \in T}$ est appelé **processus gaussien** si la combinaison linéaire finie $\sum_{i=1}^n a_i X_{t_i}$ est normalement distribuée pour tout $n \in \mathbb{N}$, $a_1, \dots, a_n \in \mathbb{R}$ et $t_1, \dots, t_n \in T$.

⁷Les terminologies « Lévy-stable » et « Lévy alpha-stable » sont également utilisées.

Exemples 3.2.8

Parmi les processus gaussiens les plus importants se trouvent :

- le processus de Wiener (non-stationnaire, mais à incréments stationnaires indépendants) ;
- le processus d’Ornstein-Uhlenbeck (stationnaire) ;
- le pont brownien (incrément non indépendants) ;
- le mouvement brownien fractionnaire (voir ci-dessous).

Définition 3.2.9

Un **mouvement brownien fractionnaire** sur $[0, T]$, noté $B_H(t)$ pour un nombre réel H donné, est un processus gaussien tel que

- i) $B_H(0) = 0$;
- ii) $\mathbb{E}[B_H(t)] = 0$ pour tout $t \in [0, T]$;
- iii) $\mathbb{E}[B_H(t)B_H(s)] = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H})$ pour tous $s, t \in [0, T]$.

Le nombre H est appelé **exposant d’autoaffinité** et respecte la contrainte $0 < H < 1$. La dernière condition dicte la fonction de covariance de la distribution.

Remarque 3.2.10 - Dans le cas particulier où $H = \frac{1}{2}$, nous obtenons un mouvement brownien [43][51].

Les mathématiques fractales s’intéressent à l’étude plus générale des fonctions $X(t)$, sur $[0, T]$, telles que $X(ct) = M(c)X(t)$ où X et M sont des variables aléatoires indépendantes. Cette relation peut être réécrite sous la forme $X(ct) = c^{H(c)}X(t)$ en prenant $M(c) = c^{H(c)}$. Dans ce contexte, $H(c)$ est une variable aléatoire et permet donc un éventail très large de valeurs possibles. Mandelbrot *et al.* imposent la condition supplémentaire $M(ab) = M_1(a)M_2(b)$ où M_1, M_2 sont deux copies indépendantes de M . Cette dernière condition mène à la définition ci-dessous d’un processus multifractal [41].

Définition 3.2.11

Un processus stochastique $X(t)$ est dit **multifractal** si ses incréments sont stationnaires et

qu'il remplit la condition suivante :

$$\mathbb{E}(|X(t)|^q) = c(q)t^{\tau(q)+1}, \quad \text{pour tout } t \in \mathcal{T}, q \in \mathcal{Q}, \quad (3.7)$$

où \mathcal{T} et \mathcal{Q} sont des intervalles de la droite réelle, $\tau(q)$ et $c(q)$ sont des fonctions de domaine \mathcal{Q} . De plus, il est supposé que \mathcal{T} et \mathcal{Q} sont de longueur positive et que $0 \in \mathcal{T}$, $[0, 1] \subseteq \mathcal{Q}$.

En imposant $c(0) = 1$, l'équation (3.7) nous donne $\tau(0) = -1$ lorsque $q = 0$. Un processus autoaffine d'index H est multifractal avec $\tau(q) = Hq - 1$. Ainsi, la « fractalité » du processus est définie par l'unique valeur H : c'est le modèle *unifractal*.

Problématiques et critique du modèle

Le modèle multifractal proposé par Mandelbrot est relativement nouveau et, nonobstant le fait que bon nombre de gens sont conscients de son existence, personne ne semble réellement l'utiliser : Mandelbrot lui-même semble davantage le mettre en valeur à des fins de simulation plutôt que de l'appliquer aux données financières réelles afin d'en effectuer une analyse probabiliste.

D'après Mandelbrot, son plus récent modèle, *Multifractal Model of Asset Returns* [41], est une alternative prometteuse aux modèles de type ARCH. Son grand désavantage, comme l'indique son auteur, est la disette de méthodes statistiques applicables. Il sera néanmoins intéressant de voir les études qui en seront faites au cours des prochaines décennies.

3.2.4 David X. Li et les copules

Suivant l'idée de Sharpe, il est essentiel de quantifier la dépendance existant entre deux actions. Sous l'hypothèse de normalité des incréments de données financières, le coefficient de corrélation linéaire suffisait pour décrire cette dépendance. Or, il est maintenant évident que l'hypothèse de normalité est surannée : il faut donc avoir recours à d'autres outils.

Les copules [30] ont été introduites par Sklar [54] vers la fin des années 1950 afin de déterminer la relation existant entre une loi de probabilité multidimensionnelle et ses marginales.

Définition 3.2.12

Une ***n*-copule** est une fonction $C : [0, 1]^n \rightarrow [0, 1]$ telle que

- i) $C(1, \dots, 1, a_m, 1, \dots, 1) = a_m$ pour tout $m \leq n$, $a_m \in [0, 1]$
- ii) $C(a_1, \dots, a_n) = 0$ si $a_m = 0$ pour un certain $m \leq n$
- iii) C est *n*-croissante, c'est-à-dire que pour tout $a \leq b$ dans $[0, 1]^n$ où $a = (a_1, \dots, a_n)$ et $b = (b_1, \dots, b_n)$, le volume assigné par C à la *n*-boîte $[a, b] = [a_1, b_1] \times \dots \times [a_n, b_n]$ est non-négatif.

Il s'ensuit de la définition 3.2.12 qu'une copule n'est nulle autre qu'une fonction de répartition multidimensionnelle dont le support est inclus dans $[0, 1]^n$ et telle que ses marginales sont des lois uniformes sur $[0, 1]$.

Considérons le cas simple où nous cherchons à caractériser la loi bivariée de deux actions. Cette dernière contient deux informations importantes concernant les actions : leur loi marginale respective, mais aussi la dépendance entre les deux actions sans égard à ces lois marginales. On a donc recours aux copules afin de décrire de façon complète et unique la dépendance existant entre les deux actions indépendamment de toute information indésirable provenant des lois marginales.

Théorème 3.2.13 (Théorème de Sklar)

Étant donnée une loi de probabilité *n*-dimensionnelle F de lois marginales F_1, \dots, F_n , il existe une *n*-copule C tel que

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad \text{pour tout } (x_1, \dots, x_n) \in [0, 1]^n. \quad (3.8)$$

Si de plus F est continue, alors C est unique.

En pratique, on détermine une *n*-copule associée à une loi de probabilité *n*-dimensionnelle F

de lois marginales F_1, \dots, F_n de la façon suivante :

$$C(u_1, \dots, u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)) \quad \text{pour } (u_1, \dots, u_n) \in [0, 1]^n. \quad (3.9)$$

C'est cette dernière idée qui a été reprise par David X. Li en 2000 [31]. Dénotons par Φ_Σ la fonction de répartition de la loi normale bivariée de matrice de corrélation Σ et par Φ la fonction de répartition de la loi normale unidimensionnelle. Le n-copule gaussien de matrice de corrélation Σ est alors donné par :

$$C_{\Sigma,n}(u_1, \dots, u_n) = \Phi_{\Sigma,n}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n)). \quad (3.10)$$

La formulation proposée par Li est la suivante.

Soient A et B deux crédits, T_A et T_B leur temps de survie respectif. Alors :

$$\mathbb{P}[T_A < 1, T_B < 1] = \Phi_{\Sigma,2}(\Phi^{-1}(F_A(1)), \Phi^{-1}(F_B(1))) \quad (3.11)$$

où F_A et F_B sont les lois respectives des temps de survie T_A et T_B .

Suite à l'effondrement du marché boursier en 2008, le monde économique a redéfini l'ignominie de l'humanité en cherchant à identifier un unique coupable : c'est à Li et à sa formule (3.11) qu'a été décerné ce déshonneur. Dans plusieurs journaux à travers le monde et sur plusieurs sites Internet, nous pouvions lire des titres d'actualité tels « *Recipe for Disaster : The Formula That Killed Wall Street* »[50] et « *Meet the man whose big idea felled Wall Street* »[29], voire même « *The Math Equation that Destroyed the World* ». Bien que les grands quotidiens et les sites de renommée utilisaient de telles formulations afin de dénoncer de façon ironique le caractère déraisonnable de ces accusations, plusieurs autres les avançaient sérieusement. Le professeur Harry Panjer, mentor de David X. Li, compare cette lâcheté à l'affirmation voulant qu'Einstein ait été responsable de la tragédie d'Hiroshima. « En statistique, nous avons un proverbe », continue Panjer : « tous les modèles sont faux, mais quelques-uns sont utiles. [David X. Li] a fourni quelque chose, un ensemble d'outils, pour les analystes financiers ; ils en ont pris une petite partie et l'ont utilisée d'une manière

qu'il n'avait jamais prévue. » Peu de gens ont la décence de mentionner que Li a lui-même déjà exprimé sa crainte par rapport au fait que le monde prenne son modèle trop au sérieux.

Bien que Li ne soit en rien responsable de l'effondrement du marché boursier en 2008, sa formule a bel et bien été mal utilisée dans le monde de l'économie. Néanmoins, il faut reconnaître que l'hypothèse de normalité utilisée dans le modèle de Li est inadéquate.

CHAPITRE 4

Modèle LPPL

De tous les modèles disponibles — en particulier celui de Mandelbrot, qui présente un intérêt particulier —, nous avons décidé de concentrer nos efforts sur les travaux de Sornette *et al.* Mais pourquoi ? D’abord, Sornette est titulaire de la *Chair of Entrepreneurial Risks* à l’ETH Zürich (Eidgenössische Technische Hochschule Zürich) et Directeur du *Financial Crisis Observatory* à cette même institution. Une certaine notoriété règne autour du nom de Sornette depuis quelques années dans le domaine de l’économie : les assertions faites par Sornette à l’égard de son modèle ne souffrent d’ailleurs en aucune façon de modestie. Malgré cela, ces ouvrages sont cités fréquemment sans jamais que quiconque ne semble s’attarder à leur validité. Il était donc très intéressant et opportun de s’adonner à une étude plus approfondie des travaux de Sornette, Johansen et Ledoit. Secondement, le modèle proposé par Sornette offre deux avantages désirables par rapport à d’autres modèles comme celui de Mandelbrot : son application aux données réelles d’un marché financier donné est (en théorie) plutôt simple et fournit une approximation de la date prévue du krach.

4.1 D'où vient le modèle

Proposé pour la première fois par Sornette, Johansen et Bouchaud en 1996 [60], le modèle LPPL, ou **log-periodic power law**¹, a subi quelques rectifications au cours des années. [58][24][59][66] L'idée fut empruntée par Sornette *et al.* au monde de la physique : Sornette a consacré plusieurs années à étudier les événements critiques dans des systèmes naturels tels les tremblements de terre et dans des systèmes complexes. En accord avec les observations des phénomènes critiques en physique statistique, les données financières se comporteraient selon une loi de puissance décorée par des oscillations log-périodiques en temps de crise, c'est-à-dire précédant un krach ; Sornette rattache notamment cette particularité aux fractales.

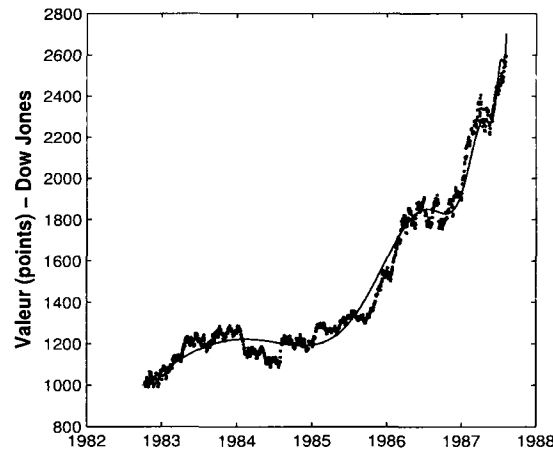


Figure 4.1 : Oscillations log-périodiques précédant un krach

D'après les travaux de Sornette *et al.*, les prix d'un marché en présence d'une bulle peuvent être fidèlement décrits par la formule suivante :

$$\ln p(t) = A + B(t_c - t)^\beta \{1 + C \cos(\omega \ln(t_c - t) + \phi)\}. \quad (4.1)$$

C'est la formule à laquelle nous ferons allusion, plus souvent qu'autrement, en discutant du *modèle LPPL*.

¹Une traduction possible serait « loi de puissance à oscillations log-périodiques » ou tout simplement « loi de puissance log-périodique ».

En posant $y(t) = \ln p(t)$, l'équation (4.1) correspond au développement de Fourier du premier ordre de la loi de puissance qui est solution de « l'équation de renormalisation »

$$\frac{dy(t)}{d \ln(t_c - t)} = z + i\omega$$

autour du « point critique » t_c . [28][26]

4.1.1 Obtention de la formule

Bien que la description du modèle (et son nom) repose sur l'apparence visuelle du graphique des fluctuations de prix d'un marché en présence d'une bulle, le développement théorique du modèle repose étonnamment sur un concept beaucoup plus intrinsèque : l'**imitation locale auto-renforçante**².

Les deux hypothèses fondamentales du modèle sont notamment les suivantes [55] :

- i) Un krach peut être causé par de l'imitation locale auto-renforçante entre les individus. Ce processus d'imitation cause l'apparition d'une bulle économique. Si la tendance des individus à imiter leurs consocuteurs et confrères augmente jusqu'à un certain point, appelé point « critique », plusieurs individus peuvent tenter de vendre simultanément, causant ainsi un krach.
- ii) Puisque le krach n'est pas une issue déterministe certaine de la bulle, il demeure rationnel pour les investisseurs de rester dans le marché en échange du taux de croissance plus élevé que leur offre la bulle. Il existe notamment une probabilité finie non nulle que la bulle se termine doucement, sans krach.

Sornette, Johansen et Bouchaud ont choisi d'exprimer leur perception des marchés boursiers en dictant les règles de décision internes du système. Grossièrement, ils stipulent que tout investisseur effectue des choix rationnels en fonction de l'information publiquement disponible et de son expérience personnelle, mais se base également sur les choix de son entourage

²Traduction possiblement mauvaise de « local self-reinforcing imitation ».

pour effectuer les siens : la décision d'un individu est donc influencée par celle des autres, jusqu'à un certain point. Afin de modéliser ce phénomène, Sornette *et al.* utilisent le *réseau hiérarchique en diamant*, sujet traité (mais non introduit) notamment dans la thèse doctorale de Johansen. Les systèmes hiérarchiques sont des systèmes itératifs qui présentent une autosimilarité : c'est l'équivalent des fractales en temps discret !

Dynamique du prix et taux de risque

Nous cherchons à modéliser la dynamique d'un marché financier précédant un krach. Considérons le *taux de risque*, que nous noterons par $h(t)$, soit la probabilité par unité de temps qu'un krach survienne dans les instants qui suivent étant donné qu'il ne soit pas déjà survenu. Soit $\kappa \in [0, 1]$, la proportion par laquelle le prix doit chuter au cours du krach, et soient finalement $p(t)$, le prix au temps t , ainsi que j , un processus de saut qui prend la valeur *un* après le krach et *zéro* avant. Notons par $\mathbb{E}_t[\cdot]$ l'espérance conditionnelle étant donnée l'information jusqu'au temps t . Sornette *et al.* proposent de modéliser la dynamique du marché selon la théorie des anticipations rationnelles, de sorte qu'une hausse de prix compense (tout juste) pour le risque associé au krach. Ainsi,

$$dp(t) = \mu(t)p(t) dt - \kappa p(t) dj \quad (4.2)$$

où $\mu(t)$ est choisi de sorte que le processus vérifie la propriété de martingale, c'est-à-dire

$$\begin{aligned} 0 &= \mathbb{E}_t[dp(t)] \\ &= \mu(t)p(t) dt - \kappa p(t)\mathbb{E}_t[dj]. \end{aligned} \quad (4.3)$$

De la définition de $h(t)$, il est clair que $\mathbb{E}_t[dj] = h(t) dt$, d'où

$$0 = \mu(t)p(t) dt - \kappa p(t)h(t) dt. \quad (4.4)$$

De l'équation (4.4), nous concluons que $\mu(t) = \kappa h(t)$. Or $dj = 0$ à tout moment précédant le krach. Ainsi, les équations (4.4) et (4.2) nous donnent

$$dp(t) = \kappa h(t)p(t) dt, \quad (4.5)$$

d'où

$$\frac{dp(t)}{p(t)} = \kappa h(t) dt \quad (4.6)$$

et ainsi

$$\ln p(t) = \ln p(t_0) + \kappa \int_{t_0}^t h(t) dt, \quad (4.7)$$

pour tout $t > t_0$ précédant le moment du krach.

Réseau hiérarchique en diamant

Considérons un système d'individus indexés par $i = 1, \dots, I$ et notons par $N(i)$ l'ensemble des individus directement connectés à l'individu i par l'entremise d'un graphe à déterminer. Supposons que chaque individu ne puisse se trouver que dans un seul de deux états, plus spécifiquement $s_i \in \{-1, +1\}$ où s_i correspond à l'état de l'individu i (à titre d'exemple, les deux états pourraient être « maintenir la position actuelle » et « inverser la position actuelle »). C'est le modèle de Potts [64].

Afin de modéliser l'interaction entre les individus, Sornette *et al.* choisissent de déterminer l'état de l'individu i à l'aide de la formule

$$s_i = \text{sgn} \left(K \sum_{j \in N(i)} s_j + \sigma \epsilon_i \right) \quad (4.8)$$

où $\text{sgn}(\cdot) \in \{-1, +1\}$ correspond au signe algébrique³, K et σ sont des constantes positives et ϵ_i sont des variables aléatoires indépendantes et identiquement distribuées selon une loi normale de moyenne *zéro* et de variance *un*. La constante K représente la tendance de l'individu i à imiter les autres, alors que la constante σ dicte la tendance de l'individu à opter pour un comportement idiosyncrasique ; c'est en fait le rapport de ces deux valeurs qui déterminera la tendance globale de l'investisseur.

Or, dans un système où les interactions entre les divers individus sont modifiées à chaque itération, il est inconcevable de croire que la valeur de K demeure constante : considérons

³Nous considérons que $\text{sgn}(x) = +1$ lorsque $x = 0$.

plutôt $K(t)$, une fonction à valeurs réelles qui peut varier dans le temps, soit pour $t \geq 0$. Aucune supposition n'est faite sur la dynamique de K , à savoir s'il s'agit d'une fonction donnée, d'une variable aléatoire ou d'un processus stochastique. Par contre, Sornette *et al.* effectuent une analogie avec un modèle utilisé dans le contexte des phénomènes critiques, notamment le modèle de Lenz-Ising bidimensionnel. Dans ce modèle, il existe une valeur critique K_c qui détermine la dynamique globale du système : lorsque $K < K_c$, le chaos règne. Toutefois, lorsque K croît et s'approche de la valeur K_c , l'ordre s'installe dans le système, c'est-à-dire qu'une légère perturbation du système affecterait grandement son issue globale. Cet ordre se traduit par le fait que les individus sont très influencés par les choix de leurs voisins et qu'en conséquence les individus rattachés de près dans le réseau tendent à se grouper suite à un commun accord concernant le choix à effectuer.

Sornette *et al.* s'intéressent particulièrement à ce qu'ils appellent la *susceptibilité* du système, soit la « probabilité » qu'un gros groupe d'individus se trouve soudainement en accord concernant leur choix. Formellement, reprenons l'équation (4.8) et ajoutons-y un terme d'influence globale G :

$$s_i = \text{sgn} \left(K \sum_{j \in N(i)} s_j + \sigma \epsilon_i + G \right). \quad (4.9)$$

Si $G > 0$, l'état des individus aura tendance à être $+1$; si $G < 0$, l'état des individus aura tendance à être -1 . Dans le cas précis où $G = 0$, nous retrouvons l'équation (4.8).

Définissons l'état moyen M du système :

$$M = \frac{1}{I} \sum_{i=1}^I s_i. \quad (4.10)$$

Sornette *et al.* argumentent que $\mathbb{E}(M) = 0$ lorsque l'influence globale $G = 0$ et que $\mathbb{E}(M) \times G \geq 0$ sinon, soit que $\mathbb{E}(M)$ et G ont le même signe algébrique. Ils définissent alors la *susceptibilité du système* comme étant

$$\chi = \left. \frac{d(\mathbb{E}[M])}{dG} \right|_{G=0}. \quad (4.11)$$

Nous pouvons penser à cette mesure comme étant la « susceptibilité » du système à opter

pour une tendance globale lorsqu'une telle influence y est exercée : comment (et à quelle « vitesse ») réagit le système lorsqu'on augmente l'influence globale ?

Dans le cas du modèle de Lenz-Ising, la susceptibilité du système tend vers l'infini au point critique K_c suivant une loi de puissance, notamment :

$$\chi = A(K_c - K)^{-\lambda}$$

où A est une constante positive et $\lambda > 0$ est appelé *exposant critique* de la susceptibilité.

Dans un contexte plus général, la structure du graphe choisi pour modéliser l'interaction entre les individus affecte la susceptibilité du système. Ainsi, Sornette *et al.* optent pour un modèle dont la solution χ est bien connue : le *réseau hiérarchique en diamant*. Commençons par un simple graphe comportant deux individus reliés par une arête. Remplaçons à l'étape suivante l'unique arête par un « diamant » (ou losange) où les deux individus de l'étape précédente occupent deux sommets opposés et les deux nouveaux sommets sont occupés par deux nouveaux individus. À chaque étape postérieure, nous remplaçons toutes les arêtes par un tel diamant. Les trois premières itérations sont illustrées dans la figure 4.2.

Nous obtenons ainsi un réseau dans lequel les plus récents investisseurs sont voisins de seulement deux individus, alors que les deux investisseurs initiaux ont beaucoup de voisins ; les autres se situent quelque part dans le milieu. Cette approche, bien que réductrice, incorpore

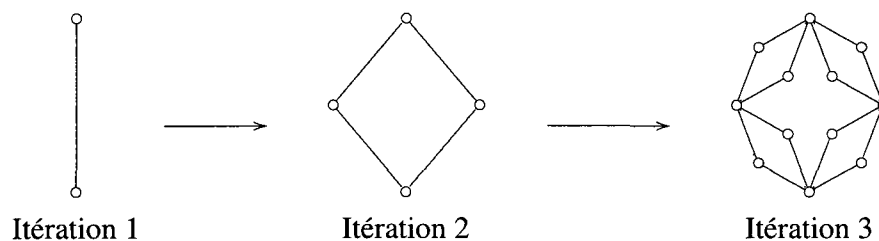


Figure 4.2 : Réseau hiérarchique en diamant (3 itérations)

deux concepts fondamentaux : les investisseurs sont influencés par les choix d'un certain nombre d'autres individus et ce nombre diffère d'un investisseur à l'autre. La susceptibilité d'un système dans lequel le réseau est ainsi défini et pour lequel l'état des investisseurs est donné par l'équation (4.8) a été établie par Derrida *et al.* [14] en 1983. Cette solution agit comme les phénomènes critiques apparaissant dans la nature : il existe une valeur critique K_c de K qui régit le comportement du système et une loi de puissance dicte le comportement global. L'exposant qui intervient dans la loi de puissance est, toutefois, un nombre complexe. Le développement de Taylor du premier ordre de la solution générale est donné par :

$$\chi \approx \text{Re}(A_0(K_c - K)^{-\lambda} + A_1(K_c - K)^{-\lambda+i\omega} + \dots) \quad (4.12)$$

$$\approx A'_0(K_c - K)^{-\lambda} + A'_1(K_c - K)^{-\lambda} \cos[\omega \ln(K_c - K) + \psi] + \dots \quad (4.13)$$

où $A_0, A_1, \lambda, \omega$ et ψ sont des valeurs réelles, et $\text{Re}[\cdot]$ désigne la partie réelle d'un nombre complexe. Pour $K < K_c$, c'est le chaos : la susceptibilité est faible, les investisseurs sont difficilement influençables et l'imitation se propage uniquement à des voisins rapprochés. Lorsque K s'approche de K_c , un état d'équilibre s'instaure : la susceptibilité croît brusquement et les amas d'investisseurs en accord grossissent. Les investisseurs étant quasiment tous en accord, une décision commune est prise et un krach s'ensuit. La loi de puissance présente dans l'équation (4.13) est ainsi « corrigée » (ou décorée) par une oscillation dite « log-périodique » dont la fréquence explose lorsque K tend vers K_c .

Afin d'appliquer le développement de Taylor autour du point critique K_c , nous devons supposer que K est lisse. Notons par t_c le premier moment où $K(t_c) = K_c$. Sornette *et al.* ont alors recours à l'approximation suivante qu'ils ne justifient pas :

$$K_c - K \approx c_0(t_c - t) \quad (4.14)$$

où c_0 est une constante et $t < t_c$. À l'aide de cette approximation, ils **postulent** que le taux de risque est de la même forme que χ , soit :

$$h(t) \approx B_0(t_c - t)^{-\alpha} + B_1(t_c - t)^{-\alpha} \cos[\omega \ln(t_c - t) + \psi']. \quad (4.15)$$

Ainsi, de l'équation (4.7) nous tirons :

$$\begin{aligned}\ln p(t) &= \ln p(t_0) + \kappa \int_{t_0}^t (B_0(t_c - t')^{-\alpha} + B_1(t_c - t')^{-\alpha} \cos[\omega \ln(t_c - t) + \psi']) dt' \\ &= A + \kappa \int_{t_0}^t B'(t_c - t')^{-\alpha} \{1 + C' \cos[\omega \ln(t_c - t') + \psi']\} dt'\end{aligned}\quad (4.16)$$

où $A := \ln p(t_0)$, $B' := B_0$ et $C' := \frac{B_1}{B_0}$.

Afin d'évaluer l'intégrale dans la partie de droite de l'équation (4.16), posons $\beta := 1 - \alpha$ et $\psi''(t) := \omega \ln(t_c - t) + \psi'$. L'équation (4.16) devient :

$$\ln p(t) = A + \kappa \int_{t_0}^t B'(t_c - t')^{\beta-1} \{1 + C' \cos \psi''(t')\} dt'. \quad (4.17)$$

Une double intégration par parties nous donne

$$\int_{t_0}^t (t_c - t')^{\beta-1} \cos \psi'' dt' = \frac{-(t_c - t)^{\beta}}{\omega^2 + \beta^2} \{ \omega \sin \psi''(t) + \beta \cos \psi''(t) \}, \quad (4.18)$$

d'où

$$[\ln p(t')]_{t'=t_0}^{t'=t} = \kappa \left[\frac{-B'}{\beta} (t_c - t')^{\beta} - \frac{B'C'}{\omega^2 + \beta^2} (t_c - t')^{\beta} \{ \omega \sin \psi''(t') + \beta \cos \psi''(t') \} \right]_{t'=t_0}^{t'=t}, \quad (4.19)$$

ou encore

$$\begin{aligned}\ln p(t) &= \kappa \left[\frac{-B'}{\beta} \{ (t_c - t)^{\beta} - (t_c - t_0)^{\beta} \} \right. \\ &\quad \left. - \frac{B'C'}{\omega^2 + \beta^2} \{ (t_c - t)^{\beta} (\omega \sin \psi''(t) + \beta \cos \psi''(t)) \right. \\ &\quad \left. - (t_c - t_0)^{\beta} (\omega \sin \psi''(t_0) + \beta \cos \psi''(t_0)) \} \right] + \ln p(t_0).\end{aligned}\quad (4.20)$$

Or, lorsque $t = t_c$, l'équation (4.20) nous donne :

$$\ln p(t_c) = \kappa \left[\frac{B'}{\beta} (t_c - t_0)^{\beta} + \frac{B'C'}{\omega^2 + \beta^2} \{ (t_c - t_0)^{\beta} (\omega \sin \psi''(t_0) + \beta \cos \psi''(t_0)) \} \right] + \ln p(t_0),$$

d'où

$$\ln p(t) = \ln p(t_c) + \left[\frac{-\kappa B'}{\beta} (t_c - t)^{\beta} - \frac{\kappa B'C'}{\omega^2 + \beta^2} (t_c - t)^{\beta} \{ \omega \sin \psi''(t) + \beta \cos \psi''(t) \} \right]. \quad (4.21)$$

En posant $A := \ln p(t_c)$, $B := \frac{-\kappa B'}{\beta}$ et $C'' := \frac{\beta C'}{\omega^2 + \beta^2}$, nous avons :

$$\ln p(t) = A + B(t_c - t)^{\beta} \{ 1 + C'' (\omega \sin \psi''(t) + \beta \cos \psi''(t)) \}. \quad (4.22)$$

Posons $C := C'' \frac{\sqrt{\omega^2 + \beta^2}}{\omega}$ et soit $\phi' \in [-\pi, \pi[$ tel que

- i) $\cos \phi' = \frac{C''\beta}{C}$
- ii) $\sin \phi' = \frac{C''\omega}{C}$
- iii) $\tan \phi' = \frac{\beta}{\omega}$.

À l'aide de l'identité trigonométrique $\cos(\psi + \phi') = \cos \psi \cos \phi' - \sin \psi \sin \phi'$, nous obtenons

$$C''(\omega \sin \psi''(t) + \beta \cos \psi''(t)) = C \cos(\psi''(t) + \phi'). \quad (4.23)$$

Ainsi,

$$\begin{aligned} \ln p(t) &= A + B(t_c - t)^\beta \{ 1 + C \cos(\psi''(t) + \phi') \} \\ &= A + B(t_c - t)^\beta \{ 1 + C \cos(\omega \ln(t_c - t) + \psi' + \phi') \} \\ &= A + B(t_c - t)^\beta \{ 1 + C \cos(\omega \ln(t_c - t) + \phi) \} \end{aligned}$$

où $\phi = \psi' + \phi'$ et nous retrouvons l'équation (4.1).

La formule (4.1) comporte sept paramètres dont quatre sont non-linéaires (β , ω , t_c , ϕ) et trois sont linéaires (A , B , C). En voici un bref sommaire :

Type	Paramètre	Description
(non-linéaire)	β	Exposant de la loi de puissance sous-jacente, mais laquelle exactement demeure nébuleux.
	ω	Fréquence des fluctuations pendant la bulle.
	t_c	Temps critique, c'est-à-dire le moment du krach anticipé par le modèle.
	ϕ	Paramètre de déphasage.
(linéaire)	A	Prix théorique si la bulle se rend jusqu'au temps t_c .
	B	Augmentation théorique en y_t sur l'unité de temps précédant le krach si C était près de <i>zéro</i> .
	C	Amplitude (proportion) des fluctuations autour de la croissance exponentielle.

Figure 4.3 : Paramètres LPPL

4.2 Tests empiriques

À ce jour, très peu de chercheurs ont effectué une étude empirique sérieuse du modèle LPPL appliqué à des données financières. En effet, les seules personnes semblant avoir effectué une étude complète du sujet sont les auteurs eux-mêmes, qui tentent à tout prix de faire prévaloir leur modèle. Leurs études s'étalent sur plusieurs années, suivant ainsi l'évolution de leurs travaux et du modèle en question.

Chang et Feigenbaum ont analysé quelques aspects plus théoriques du modèle LPPL à l'aide de méthodes bayésiennes [12], mais leurs études, bien que très intéressantes, ne sont en rien complètes : elles se concentrent souvent sur un unique krach restreint à un seul marché.

4.2.1 Les sept paramètres et leurs contraintes

L'application du modèle LPPL à des données financières consiste à déterminer les valeurs des paramètres qui minimisent la différence entre le modèle et les données empiriques, au sens des moindres carrés. Or, plutôt que d'établir des critères sur les résidus afin de savoir si l'ajustement du modèle aux données est convenable — et donc qu'un krach est probable, voire imminent —, Sornette *et al.* ont préféré effectuer le meilleur ajustement possible du modèle aux divers krachs connus de plusieurs marchés et d'en tirer des conclusions. Ils ont établi des intervalles, basés sur les données historiques, dans lesquels doivent se trouver les divers paramètres afin que l'ajustement indique la présence d'un krach imminent. Voici les conditions qu'ils ont établies sur les divers paramètres :

La condition $\phi \in [0, 2\pi]$ n'est en fait pas une contrainte, puisqu'il s'agit d'une variable de déphasage de la fonction cosinus ; il suffit donc d'ajouter un multiple de 2π à ϕ afin que la condition soit respectée. La condition $t_c > 0$, en pratique, devient $t_c > t_{max}$ où t_{max} est la date la plus tardive des données utilisées pour l'ajustement. Le paramètre linéaire C , quant à lui, n'est contraint d'aucune manière.

Type	Paramètre et contraintes
(non-linéaires)	$0.33 - 0.18 \leq \beta \leq 0.33 + 0.18$
	$6.36 - 1.56 \leq \omega \leq 6.36 + 1.56$
	$t_c > 0$
	$\phi \in [0, 2\pi]$
(linéaires)	$A > 0$
	$B < 0$
	$C \in \mathbb{R}$

Tableau 4.1 : Contraintes sur les paramètres LPPL

Des questions naturelles se présentent : si les intervalles établis utilisent l'ensemble des données historiques, est-ce que cela signifie qu'absolument tous les krachs survenus jusqu'à présent respectent ces conditions ? Et si oui, doit-on s'attendre à ce que tous les krachs à venir les respectent également, ou alors ces contraintes devront-elles éventuellement être modifiées suivant l'évolution des marchés ?

4.2.2 Estimation des paramètres

La constatation la plus frappante en essayant d'appliquer le modèle LPPL à des données financières est la difficulté d'estimation des paramètres. En effet, il n'est pas du tout facile de déterminer les valeurs « optimales » d'un vecteur de paramètres qui minimisent une fonction non-linéaire. Même en présence de très bons algorithmes, la convergence n'est que rarement garantie. Et si c'est le cas, rien ne nous assure que l'algorithme aura convergé vers un minimum absolu plutôt qu'un minimum local : c'est la raison d'être de la recherche opérationnelle.

Paramètres linéaires et non-linéaires

Rappelons la formule du modèle LPPL, soit l'équation (4.1) :

$$\ln p(t) = A + B(t_c - t)^\beta \{1 + C \cos(\omega \ln(t_c - t) + \phi)\}.$$

Les paramètres A , B et C sont linéaires alors que β , ω , t_c et ϕ sont non-linéaires. Nous cherchons le vecteur $(\beta, \omega, t_c, \phi, A, B, C)$ qui minimisera, au sens des moindres carrés, la différence entre les données observées et les données prédites par l'équation ci-dessus. Notons par y_i la $i^{\text{ème}}$ donnée empirique (prix) disponible (chronologiquement) et par t_i le temps correspondant à cette donnée.

Afin de simplifier les calculs dans cette section, réécrivons la formule (4.1) sous une forme équivalente, notamment :

$$\ln p(t) = A + B(t_c - t)^\beta + C'(t_c - t)^\beta \cos(\omega \ln(t_c - t) + \phi), \quad (4.24)$$

où $C' := BC$. En posant $f(t) := (t_c - t)^\beta$ et $g(t) = (t_c - t)^\beta \cos(\omega \ln(t_c - t) + \phi)$, et en utilisant les notations $y_i := y(t_i)$, $f_i := f(t_i)$ et $g_i := g(t_i)$, nous obtenons :

$$\hat{y}_i = A + Bf_i + C'g_i, \quad (4.25)$$

où \hat{y}_i est l'approximation de y_i offerte par l'ajustement du modèle aux données empiriques.

Afin de simplifier la détermination des paramètres « optimaux », l'astuce de Sornette *et al.* consiste à exprimer les paramètres linéaires en fonction des paramètres non-linéaires ; l'étape d'optimisation peut alors se faire sur quatre paramètres plutôt que sept. Minimiser la différence entre les données empiriques et les valeurs prédites par le modèle au sens des moindres carrés nous donne :

$$SE = \sum_{i=1}^N (y_i - \hat{y}_i)^2 = \sum_{i=1}^N (y_i - A - Bf_i - C'g_i)^2, \quad (4.26)$$

où N est le nombre total de données empiriques utilisées pour l'ajustement. Afin d'obtenir les valeurs de A , B et C' qui minimisent SE , nous dérivons par rapport à chacun de ces trois

paramètres et nous égalons chacune de ces dérivées à zéro ; la somme de carrés nous assure ainsi des minimums globaux. Nous obtenons donc :

$$\begin{aligned}\frac{\partial SE}{\partial A} &= -2 \sum_{i=1}^N (y_i - A - Bf_i - C'g_i) \\ &= -2 \left(\sum_{i=1}^N y_i - AN - B \sum_{i=1}^N f_i - C' \sum_{i=1}^N g_i \right) = 0,\end{aligned}$$

d'où

$$\sum_{i=1}^N y_i = AN + B \sum_{i=1}^N f_i + C' \sum_{i=1}^N g_i. \quad (4.27)$$

De même,

$$\begin{aligned}\frac{\partial SE}{\partial B} &= -2 \sum_{i=1}^N f_i (y_i - A - Bf_i - C'g_i) \\ &= -2 \left(\sum_{i=1}^N y_i f_i - A \sum_{i=1}^N f_i - B \sum_{i=1}^N f_i^2 - C' \sum_{i=1}^N f_i g_i \right) = 0,\end{aligned}$$

d'où

$$\sum_{i=1}^N y_i f_i = A \sum_{i=1}^N f_i + B \sum_{i=1}^N f_i^2 + C' \sum_{i=1}^N f_i g_i. \quad (4.28)$$

Finalement,

$$\begin{aligned}\frac{\partial SE}{\partial C'} &= -2 \sum_{i=1}^N g_i (y_i - A - Bf_i - C'g_i) \\ &= -2 \left(\sum_{i=1}^N y_i g_i - A \sum_{i=1}^N g_i - B \sum_{i=1}^N f_i g_i - C' \sum_{i=1}^N g_i^2 \right) = 0,\end{aligned}$$

d'où

$$\sum_{i=1}^N y_i g_i = A \sum_{i=1}^N g_i + B \sum_{i=1}^N f_i g_i + C' \sum_{i=1}^N g_i^2. \quad (4.29)$$

Ainsi, les formules (4.27), (4.28) et (4.29) nous fournissent un système d'équations linéaires qu'il suffit de résoudre pour déterminer les valeurs de A , B , C' à partir de celles des quatre paramètres non-linéaires cachés dans f_i et g_i :

$$\begin{pmatrix} N & \sum f_i & \sum g_i \\ \sum f_i & \sum f_i^2 & \sum f_i g_i \\ \sum g_i & \sum f_i g_i & \sum g_i^2 \end{pmatrix} \times \begin{pmatrix} A \\ B \\ C' \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum y_i f_i \\ \sum y_i g_i \end{pmatrix}. \quad (4.30)$$

Algorithme d'optimisation non-linéaire

Afin de déterminer le vecteur $(\beta, \omega, t_c, \phi)$ qui minimise l'équation (4.26), nous devons faire appel à un algorithme d'optimisation non-linéaire. La phase d'optimisation des paramètres dans les travaux de Sornette *et al.* a été effectuée de la manière suivante :

- (i) Établir dix (10) vecteurs $(\beta, \omega, t_c, \phi)$ à être utilisés comme solution initiale pour la phase d'optimisation non-linéaire.
- (ii) Pour chacun des dix (10) vecteurs initiaux, procéder à l'optimisation non-linéaire.
- (iii) Conserver la solution minimisant l'ajustement du modèle au sens des moindres carrés.

Examinons davantage chacune de ces étapes.

Étape (i) Afin de déterminer les vecteurs initiaux, Sornette *et al.* font appel à l'algorithme Tabu [20] ; il s'agit d'une métaheuristique de type « recherche locale ». Grossièrement, l'algorithme consiste à observer aléatoirement un certain nombre de points autour de la solution actuelle afin de vérifier si ceux-ci offrent une meilleure solution. À chaque itération, l'unique meilleure solution (excluant la précédente) est conservée et une boule topologique est construite autour de la solution précédente afin d'éviter d'y revenir pour un certain temps. Toutefois, étant de type « recherche locale », il est aisé de se demander comment déterminer les points initiaux à fournir à l'algorithme ; Sornette *et al.* ne fournissent pas énormément de détails à cet effet. L'approche la plus intuitive consiste à effectuer un partitionnement du domaine admissible de chacun des paramètres et de fournir chacune des combinaisons possibles comme vecteur initial. Un problème toutefois : le nombre de combinaisons croît très rapidement ! De plus, le fait de se restreindre au domaine admissible de chacun des paramètres pourrait nous empêcher de trouver une solution inadmissible qui minimiserait davantage la fonction objectif, résultant ainsi en un très indésirable faux positif. Dans les rares occasions où les auteurs décrivent leur processus d'optimisation [24], nous apprenons que leur application de l'algorithme Tabu est effectuée en deux étapes : la première consiste à appliquer l'algorithme uniquement aux paramètres β et ω en fixant les autres paramètres

non-linéaires. Toutes les valeurs pour lesquelles l'algorithme converge en respectant la condition $0 < \beta < 1$ sont alors réutilisées (les valeurs issues de la convergence et non les valeurs initiales) pour une seconde phase de l'algorithme en laissant les autres paramètres non-linéaires varier. Les dix (10) meilleures solutions peuvent alors être conservées pour l'étape suivante.

Étape (ii) L'algorithme utilisé par Sornette *et al.* est de type *Levenberg-Marquardt*. C'est un algorithme hybride très performant, au croisement entre la méthode de descente de gradient et l'algorithme de Gauss-Newton, qui est adapté aux problèmes de minimisation d'une fonction non-linéaire à plusieurs variables au sens des moindres carrés. Afin de déterminer les valeurs \hat{y}_i prédites par le modèle à chaque itération, les paramètres linéaires doivent être déterminés à l'aide du système linéaire défini par l'équation (4.30).

Étape (iii) À partir des dix (10) vecteurs initiaux, l'algorithme de Levenberg-Marquardt donne dix (10) solutions « optimales » pour l'ajustement au sens des moindres carrés. De ces dix solutions, celle dont les résidus affichent la plus petite somme des carrés est conservée comme unique meilleure solution. Si les paramètres issus de cette solution respectent les conditions énoncées dans le tableau 4.1, un krach est prévu par le modèle ; sinon, l'ajustement ne prévoit aucun krach.

4.2.3 Sornette *et al.*

Sornette, Bouchaud, Johansen, Ledoit et Zhou sont conjointement les personnes qui ont appliqué le plus souvent le modèle LPPL à des données empiriques ; ce sont également les gens qui ont travaillé sur l'élaboration et la rectification du modèle à travers les années. Bien que les prétentions soient nombreuses à l'égard du modèle et de l'application qu'ils en font à divers marchés, aucune publication n'offre une description précise des méthodes utilisées et l'ensemble complet des résultats obtenus n'est jamais divulgué. Fait intéressant : tous leurs articles mentionnent le succès de l'application du modèle à certaines séries financières, mais aucun document ne discute des situations où le modèle échoue ou ne prédit pas de krach.

Néanmoins, s'ils gardent toute l'information de leurs études pour eux, il n'en reste pas moins qu'ils ont vraisemblablement effectué ses études ; et il y en a énormément ! Parmi les plus intéressantes et complètes se trouve une étude portant sur les krachs du Hang Seng sur une période de plus de 30 ans, couvrant un total de sept krachs.

Malheureusement, les informations présentées dans les diverses publications des auteurs ci-haut mentionnés à travers les années sont incohérentes : d'un article à l'autre, des modifications non justifiées ou non justifiables sont apportées au modèle ou au processus d'ajustement et les critères d'admissibilité varient. Les auteurs, dont les travaux sont presque toujours issus d'un effort commun, se critiquent même entre eux [23].

4.2.4 Brée et Joseph

David S. Brée et Nathan L. Joseph, décelant diverses anomalies dans les publications de Sornette *et al.* ainsi que des problématiques dans le fondement théorique du modèle, ont entrepris de reproduire l'étude [10] du marché de Hang Seng effectuée par Sornette *et al.* Ils ont notamment identifié davantage de krachs que l'étude originale et recensé des problématiques liées à leur détection à l'aide du modèle LPPL. Parmi les problèmes recensés se trouvent : la difficulté d'identification du début de la bulle, la tendance de Sornette *et al.* à ne pas suivre leurs propres directives et l'imprécision du processus d'ajustement du modèle aux données.

Les critiques de Brée et Joseph à l'égard des travaux de Sornette *et al.* sont nombreuses. Plusieurs idées utilisées au cours des expérimentations décrites à la section 4.2.5 (ainsi que certaines présentées dans la section 4.3) ont été empruntées ou sont issues de leurs travaux. Une communication entretenue avec Brée aura d'ailleurs été essentielle à l'aboutissement des expérimentations de la section 4.2.5.

Brée critique l'utilisation des contraintes indiquées dans le tableau 4.1 comme critères ultimes d'identification d'un krach. Brée et Joseph démontrent notamment que la valeur SE de l'équation (4.26) est une fonction non-linéaire de ω [7]. Plus spécifiquement, des variations

de la valeur de ω produisent une oscillation sinusoïdale de la valeur de SE . Afin de contourner ce problème, Brée et Joseph utilisent une formulation alternative de l'équation (4.1) qui résout ce problème[9][10].

4.2.5 Expérimentations

L'idée est simple : reproduire le processus d'ajustement utilisé par Sornette *et al.* et l'appliquer à deux séries financières. La première doit comporter un krach historique afin de mettre à l'épreuve la capacité du modèle à identifier un krach et de nous permettre d'analyser la sensibilité de l'ajustement à différents facteurs. La seconde série doit comporter uniquement des données actuelles présentant une forte susceptibilité d'effondrement afin de vérifier la capacité du modèle à identifier un krach imminent en temps réel.

La première série de données financières choisie est l'indice du *Dow Jones (Industrial Average)* précédant le krach d'octobre 1987. La seconde est le prix de l'or en devise américaine (USD). Les deux séries financières utilisées correspondent aux valeurs quotidiennes de fermeture de la bourse.

Algorithme d'ajustement

Le processus d'ajustement a été implémenté le plus fidèlement possible afin d'être en accord avec les travaux de Sornette *et al.* [25] et se déroule comme suit :

A) Phase 1 - Tabu

Les domaines admissibles des paramètres non-linéaires, tels que donnés par le tableau 4.1, sont étendus dans les deux directions et partitionnés. Toutes les combinaisons possibles des points de partition sont soumises à une première phase rapide de l'algorithme Tabu. Durant cette phase, les paramètres t_c et ϕ sont fixés. Seuls les paramètres non-linéaires sont soumis à l'algorithme ; les paramètres linéaires sont déterminés analytiquement, à chaque itération, à l'aide du système (4.30).

Toutes les solutions issues de cette première étape et admettant $0 < \beta < 1$ sont retenues et soumises à une seconde phase plus rigoureuse de l'algorithme Tabu. Durant cette phase, les quatre paramètres non-linéaires peuvent varier ; les paramètres linéaires sont déterminés analytiquement.

Toutes les solutions admissibles (au sens du tableau 4.1) issues de cette seconde étape sont retenues. Le cas échéant, seules les 1000 meilleures solutions (au sens des moindres carrés) sont conservées.

B) Phase 2 - Levenberg-Marquardt

Les 10 meilleures solutions (au sens des moindres carrés) issues de la phase précédente sont sélectionnées.

Lorsque possible, une solution qui est « topologiquement loin » des 10 déjà sélectionnées est choisie parmi les 1000 meilleures de la phase précédente. Ceci est utile pour tester la sensibilité de l'ajustement aux paramètres initiaux.

Un certain nombre de vecteurs est sélectionné manuellement en inscrivant plus ou moins aléatoirement des valeurs de notre choix. Ceci est utile pour tester la sensibilité de l'ajustement aux paramètres initiaux.

Tous les vecteurs conservés sont soumis à l'algorithme de Levenberg-Marquardt tel qu'implanté par la fonction *lsqcurvefit* de MATLAB.

Sensibilité aux paramètres initiaux

Brée et al. ont étudié la sensibilité de l'ajustement du modèle à des données par rapport à des variations des vecteurs initiaux. Leurs résultats semblent suggérer que l'erreur de l'ajustement (soit la norme des résidus) varie énormément avec de *petites* variations du paramètre ω (les grosses variations n'étant d'aucun intérêt particulier)[7].

Or, les tests que nous avons effectués n'ont démontré *aucune sensibilité de l'ajustement à l'égard des vecteurs initiaux*. En effet, deux vecteurs initiaux quelconques convergent vers

deux solutions pratiquement identiques ; s'il s'agit d'un minimum local, tous les vecteurs initiaux utilisés convergent vraisemblablement vers le même. Le vecteur initial fourni ne semble donc être d'aucune importance, dans la mesure bien sûr où il permet à l'algorithme de converger.

Un exemple de l'ajustement optimal fourni par différents vecteurs initiaux est résumé dans le tableau suivant :

	β	ω	t_c	ϕ	A	B	C	\sqrt{SE}
Vecteur initial	0 33	6 36	1987 7	2 56	—	—	—	—
Solution	0 322880	5 397477	1987 647309	4 908772	3229 693983	-1407 199657	107 983166	1974 613089
Vecteur initial	0 31	5 5	1987 8	2 2	—	—	—	—
Solution	0 322880	5 397481	1987 647310	1 767175	3229 695270	-1407 200641	-107 983211	1974 613089
Vecteur initial	0 141215	5 5	1987 8	2 2	—	—	—	—
Solution	0 322879	5 397482	1987 647310	1 767173	3229 697385	-1407 202773	-107 983206	1974 613089
Vecteur initial	0 32	4 3	1987 7	1 7	—	—	—	—
Solution	0 322880	5 397482	1987 647310	1 767174	3229 695385	-1407 200728	-107 983215	1974 613089
Vecteur initial	0 36	5 1	1987 85	2 06	—	—	—	—
Solution	0 322879	5 397482	1987 647310	1 767174	3229 696859	-1407 202263	-107 983204	1974 613089
Vecteur initial	0 472061	5 675448	1987 762544	4 410207	—	—	—	—
Solution	0 322879	5 397482	1987 647310	4 908765	3229 697815	-1407 203225	107 983203	1974 613089
Vecteur initial	0 506923	5 489873	1987 941851	4 199018	—	—	—	—
Solution	0 322880	5 397481	1987 647309	4 908768	3229 695878	-1407 201326	107 983198	1974 613089
Vecteur initial	0 484277	5 571844	1987 755060	4 629051	—	—	—	—
Solution	0 322880	5 397483	1987 647310	4 908766	3229 696717	-1407 202061	107 983213	1974 613089
Vecteur initial	0 507536	5 577435	1987 797144	4 435965	—	—	—	—
Solution	0 322880	5 397482	1987 647310	4 908767	3229 695148	-1407 200463	107 983219	1974 613089
Vecteur initial	0 431026	5 753438	1987 772268	4 294078	—	—	—	—
Solution	0 322880	5 397482	1987 647310	4 908767	3229 695410	-1407 200750	107 983215	1974 613089
Vecteur initial	0 459698	5 662319	1987 755045	4 496644	—	—	—	—

Solution	0 322879	5 397483	1987 647310	4 908765	3229 696809	-1407 202147	107 983213	1974 613089
Vecteur initial	0 488215	6 546409	1987 886664	3 149921	—	—	—	—
Solution	0 322879	5 397482	1987 647310	4 908766	3229 696826	-1407 202231	107 983204	1974 613089
Vecteur initial	0 466928	5 742081	1987 748972	4 412784	—	—	—	—
Solution	0 322880	5 397481	1987 647310	4 908767	3229 695292	-1407 200657	107 983212	1974 613089
Vecteur initial	0 405420	5 686763	1987 802397	4 590217	—	—	—	—
Solution	0 322879	5 397483	1987 647310	4 908765	3229 696821	-1407 202159	107 983214	1974 613089
Vecteur initial	0 479565	5 881417	1987 793510	4 244426	—	—	—	—
Solution	0 322879	5 397483	1987 647310	4 908765	3229 697965	-1407 203373	107 983203	1974 613089
Vecteur initial	0 201437	7 458980	1988 260000	4 434690	—	—	—	—
Solution	0 322879	5 397482	1987 647310	8 050359	3229 697019	-1407 202418	-107 983205	1974 613089

Tableau 4.2 : Sensibilité aux vecteurs initiaux — Dow Jones (1987)

Étant donnés les résultats que nous avons obtenus, la première phase de l'algorithme d'ajustement suggéré par Sornette *et al.* semble donc passablement inutile, sinon peut-être pour permettre l'identification de vecteurs initiaux qui assurent la convergence de l'algorithme.

Sensibilité à la fenêtre temporelle

Lors de l'ajustement, il est primordial de fixer une fenêtre temporelle dans laquelle les données seront observées. Afin de tester la sensibilité de l'ajustement à cette fenêtre temporelle, nous avons fixé une date de fin et nous avons manuellement sélectionné les dates qui semblaient les plus susceptibles de correspondre à un début de bulle économique. Puisqu'un modèle de prédiction de krach doit être utile pour détecter un effondrement à tout moment précédant cet événement, la date de fin ne devrait être d'aucune importance ; cette dernière a donc été fixée à un mois précédant la date réelle du krach pour l'ensemble des essais.

Les tout premiers essais que nous avons effectués semblaient laisser supposer que le choix

d'une date de début pour la fenêtre temporelle était d'une importance primordiale. En effet, le choix de deux dates distinctes fournit des ajustements très différents. Le tableau suivant illustre la situation. Dans ce dernier, t_{min} indique la date de début utilisée pour l'ajustement et t_{max} correspond à la date de fin ; les autres données correspondent à des ajustements finaux. Notons ici que SE est une mesure inutile à des fins de comparaison des diverses fenêtres étant données leurs longueurs respectives.

t_{min}	t_{max}	β	ω	t_c	ϕ	A	B	C	$RMSE$
		0.616545	5.654171	1987.713792	1.423449	7.958043	-0.429719	-0.060989	0.00116820
1982.610959	1987.712329	0.612500	5.667837	1987.715985	1.401820	7.962122	-0.433473	-0.061000	0.00116893
		0.597275	5.718671	1987.734336	1.308763	7.983065	-0.449856	-0.061076	0.00116984
		0.519294	5.746202	1987.777933	1.107990	8.073222	-0.533547	-0.059952	0.00111291
1982.775342	1987.712329	0.519055	5.747856	1987.778224	4.247059	8.073650	-0.533899	0.059956	0.00111291
		0.559920	5.589207	1987.734607	4.518479	8.008968	-0.479736	0.059690	0.00110988
1984.453552	1987.712329	0.763986	9.455222	1988.084385	4.239330	8.082558	-0.422989	-0.031743	0.00091785
		0.763889	9.457520	1988.084847	4.235353	8.082806	-0.423090	-0.031740	0.00091785
1984.560109	1987.712329	0.746925	9.069042	1988.043711	4.672664	8.075673	-0.430312	-0.032670	0.00093452
1985.764384	1987.712329	0.729685	8.042313	1987.941967	5.638721	8.044090	-0.436362	-0.053973	0.00093944

Tableau 4.3 : Sensibilité à la fenêtre temporelle — Dow Jones (1987)

Les résultats finaux de l'ajustement aux données d'octobre 1987 — ainsi que l'ajustement aux prix de l'or — portent toutefois à croire que la fenêtre temporelle choisie n'est pas particulièrement importante.

Ajustement au krach d'octobre 1987

Après avoir déterminé les vecteurs initiaux à l'aide de l'algorithme Tabu, soumis des milliers de vecteurs initiaux à l'algorithme de Levenberg-Marquardt et répété le processus pour une

variété de fenêtres temporelles, une tendance se maintenait : **aucun ajustement** du modèle au logarithme des données de l'indice du Dow Jones précédant le krach d'octobre 1987 **n'offrait une solution admissible** ! En d'autres termes, selon les travaux de Sornette *et al.*, le modèle LPPL est incapable de prédire le krach d'octobre 1987, l'un des pires effondrements boursiers de tous les temps. Il suffit de comparer les valeurs de β et ω du tableau 4.3 aux contraintes du tableau 4.1 afin de constater la situation.

Toutefois, Sornette *et al.* effectuent la quasi-totalité de leurs ajustements en utilisant les données brutes (soit les prix réels plutôt que leur logarithme) ; tous les ajustements ont donc été refaits à l'aide des données brutes. Les résultats sont percutants : **absolument tous les ajustements**, nonobstant la fenêtre temporelle utilisée et les vecteurs initiaux sélectionnés, **fournissent des solutions admissibles** ! En d'autres termes, le fait de considérer les données brutes plutôt que le logarithme des données fait pivoter le verdict de « aucun krach envisagé » à « krach imminent ».

	β	ω	t_c	ϕ	A	B	C	\sqrt{SE}
Vecteur initial	0 505271	5 766494	1987 782734	1 056569	—	—	—	—
Solution	0 227049	6 011706	1987 806513	0 734653	4147 898258	-2248 136707	-109 142950	2220 970559
Vecteur initial	0 508485	5 821435	1987 783126	0 999597	—	—	—	—
Solution	0 227052	6 011690	1987 806511	0 734676	4147 862590	-2248 101625	-109 142916	2220 970560
Vecteur initial	0 507737	5 707688	1987 777260	4 279207	—	—	—	—
Solution	0 227057	6 011655	1987 806509	3 876314	4147 806745	-2248 046719	109 142813	2220 970561
Vecteur initial	0 491675	5 814731	1987 782322	4 176010	—	—	—	—
Solution	0 227058	6 011646	1987 806508	3 876324	4147 788672	-2248 028638	109 142797	2220 970562
Vecteur initial	0 503968	5 781129	1987 787874	1 082529	—	—	—	—
Solution	0 227048	6 011710	1987 806513	0 734648	4147 901249	-2248 139635	-109 142969	2220 970559
Vecteur initial	0 499709	5 753070	1987 785757	7 398255	—	—	—	—
Solution	0 227050	6 011701	1987 806513	7 017844	4147 888521	-2248 127137	-109 142940	2220 970559
Vecteur initial	0 502449	5 739031	1987 770125	1 112791	—	—	—	—

Solution	0.227068	6.011587	1987.806502	0.734813	4147.670561	-2247.912512	-109.142655	2220.970567
Vecteur initial	0.506818	5.893147	1987.789755	7.181469	—	—	—	—
Solution	0.227042	6.011751	1987.806517	7.017778	4147.974103	-2248.211308	-109.143081	2220.970558
Vecteur initial	0.502670	5.807945	1987.771482	4.188418	—	—	—	—
Solution	0.227064	6.011614	1987.806505	3.876370	4147.720402	-2247.961518	109.142724	2220.970565
Vecteur initial	0.494291	5.783253	1987.794275	1.056151	—	—	—	—
Solution	0.227092	6.011449	1987.806489	0.735004	4147.395726	-2247.642368	-109.142319	2220.970591
Vecteur initial	0.203038	6.725072	1988.131348	2.200455	—	—	—	—
Solution	0.227083	6.011558	1987.806496	3.876467	4147.505688	-2247.749277	109.142828	2220.970574

Tableau 4.4 : Ajustement au Dow Jones (1987) — Données brutes

Les meilleures solutions issues des ajustements au logarithme des données ainsi qu’aux données brutes du Dow Jones précédant le krach d’octobre 1987 sont illustrées ci-dessous.

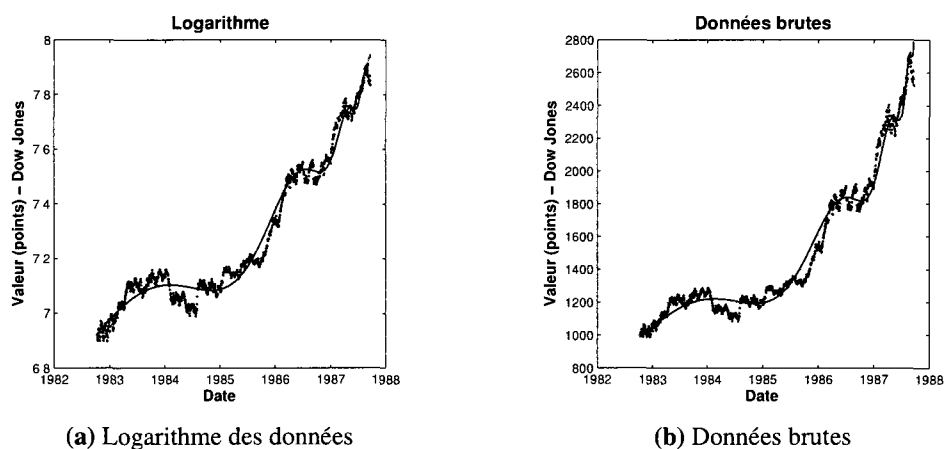


Figure 4.4 : Meilleurs ajustements au Dow Jones — Krach d’octobre 1987

Une question survient donc naturellement : les contraintes d’admissibilité établies dans le tableau 4.1 sont-elles basées uniquement sur des ajustements effectués à l’aide de données brutes ? En effet, les ajustements de Sornette *et al.* sont plus souvent qu’autrement effectués

sur des données brutes et les contraintes d'admissibilité ont été établies en fonction d'ajustements du modèle à des krachs historiques. Alors que tous les essais que nous avons effectués sur les données brutes ont fourni des valeurs de $\beta \in [.15, .51]$, les ajustements au logarithme des données ont quant à eux constamment livré des valeurs approximatives de $\beta \in [.518, .80]$.

Néanmoins, l'ajustement aux données brutes prédit un krach au 22 octobre 1987, soit seulement trois jours après la date réelle du début du krach. Ainsi, la valeur de $\beta \approx 0.32$ obtenue est très similaire à celle fournie par Sornette [56], soit $\beta \approx 0.33$. Aucune information concernant les valeurs des autres paramètres issues de l'ajustement effectué par Sornette *et al.* n'est disponible dans l'article.

Une chose est toutefois certaine : l'ajustement du modèle aux données brutes ou à leur logarithme n'offre pas les mêmes valeurs « optimales » des paramètres. L'utilisation intermittente des deux méthodes (l'ajustement aux données brutes et au logarithme) remet donc en question la validité des contraintes du tableau 4.1.

Ajustement aux prix actuels de l'or

En utilisant une fenêtre temporelle suffisamment grande — une fenêtre trop petite empêchant l'exécution de l'algorithme —, les résultats de l'ajustement du modèle LPPL aux prix de l'or (en devise américaine) ont été similaires à ceux obtenus lors de l'ajustement aux données du Dow Jones précédant le krach d'octobre 1987. Plus précisément, tous les ajustements effectués aux données brutes annoncent l'imminence d'un krach, alors que tous les ajustements effectués au logarithme des données offrent des solutions inadmissibles.

Les résultats obtenus de l'ajustement au logarithme des données sont les suivants :

	β	ω	t_c	ϕ	A	B	C	\sqrt{SE}
Vecteur initial	0.496456	6.759470	2011.412584	4.620670	—	—	—	—
Solution	0.641149	6.031579	2011.146298	6.012663	7.365340	-0.349192	0.030390	2.416205

Vecteur initial	0 506483	7 054811	2011 565442	3 880559	–	–	–	–
Solution	0 641149	6 031500	2011 146274	6 012810	7 365334	-0 349191	0 030391	2 416205
Vecteur initial	0 509961	5 813149	2011 149730	6 200011	–	–	–	–
Solution	0 641146	6 031549	2011 146295	6 012709	7 365342	-0 349195	0 030390	2 416205
Vecteur initial	0 488630	6 077329	2011 206382	2 695914	–	–	–	–
Solution	0 641145	6 031551	2011 146297	2 871108	7 365343	-0 349195	-0 030390	2 416205
Vecteur initial	0 490243	6 218768	2011 231200	2 418053	–	–	–	–
Solution	0 641146	6 031537	2011 146291	2 871137	7 365341	-0 349194	-0 030390	2 416205
Vecteur initial	0 500969	6 646033	2011 466964	1 524516	–	–	–	–
Solution	0 641148	6 031516	2011 146281	2 871184	7 365337	-0 349192	-0 030390	2 416205
Vecteur initial	0 504408	6 673097	2011 445299	4 490931	–	–	–	–
Solution	0 641148	6 031622	2011 146310	6 012585	7 365345	-0 349194	0 030390	2 416205
Vecteur initial	0 476993	6 698871	2011 465187	4 547876	–	–	–	–
Solution	0 641148	6 031515	2011 146281	6 012778	7 365337	-0 349192	0 030391	2 416205
Vecteur initial	0 472933	6 890311	2011 489846	4 239032	–	–	–	–
Solution	0 641148	6 031510	2011 146279	6 012788	7 365336	-0 349192	0 030391	2 416205
Vecteur initial	0 489818	6 076229	2011 170205	6 010060	–	–	–	–
Solution	0 641145	6 031552	2011 146297	6 012701	7 365343	-0 349195	0 030390	2 416205
Vecteur initial	0 186440	5 970107	2011 231866	-0 098513	–	–	–	–
Solution	0 641145	6 031555	2011 146300	-0 270495	7 365344	-0 349196	0 030390	2 416205
Vecteur initial	0 263123	7 887407	2011 964151	1 754966	–	–	–	–
Solution	0 641149	6 031576	2011 146297	6 012669	7 365340	-0 349192	0 030390	2 416205
Vecteur initial	0 264453	7 791482	2012 081755	4 760783	–	–	–	–
Solution	0 641148	6 031571	2011 146298	9 154264	7 365342	-0 349193	-0 030390	2 416205

Tableau 4.5 : Ajustement aux prix de l'or (USD) — Logarithme

Les résultats obtenus de l'ajustement aux données brutes sont les suivants :

	β	ω	t_c	ϕ	A	B	C	\sqrt{SE}
Vecteur initial	0 706873	5 297169	2010 705650	7 764465	–	–	–	–
Solution	0 416105	6 277380	2010 828265	6 642994	1510 134120	-475 552257	36 057594	1842 039860
Vecteur initial	0 508218	5 866903	2010 791592	0 823454	–	–	–	–
Solution	0 416105	6 277379	2010 828265	0 359811	1510 133603	-475 551833	36 057589	1842 039860
Vecteur initial	0 509011	5 961398	2010 815659	0 581935	–	–	–	–
Solution	0 416107	6 277382	2010 828262	0 359818	1510 129596	-475 548401	36 057552	1842 039860
Vecteur initial	0 505207	6 001170	2010 813543	3 794903	–	–	–	–
Solution	0 416107	6 277384	2010 828262	3 501407	1510 130210	-475 548892	-36 057559	1842 039860
Vecteur initial	0 498226	5 927288	2010 796665	0 735089	–	–	–	–
Solution	0 416106	6 277455	2010 828270	0 359719	1510 135729	-475 552592	36 057527	1842 039860
Vecteur initial	0 501179	6 149630	2010 830769	3 530531	–	–	–	–
Solution	0 416107	6 277383	2010 828262	3 501409	1510 129791	-475 548557	-36 057554	1842 039860
Vecteur initial	0 504268	6 049905	2010 836710	3 652219	–	–	–	–
Solution	0 416105	6 277465	2010 828271	3 501298	1510 137033	-475 553558	-36 057531	1842 039860
Vecteur initial	0 494681	6 072749	2010 814648	3 682056	–	–	–	–
Solution	0 416105	6 277379	2010 828265	3 501404	1510 133608	-475 551839	-36 057589	1842 039860
Vecteur initial	0 492432	5 994405	2010 805286	0 608703	–	–	–	–
Solution	0 416107	6 277425	2010 828266	0 359761	1510 131890	-475 549744	36 057513	1842 039860
Vecteur initial	0 505829	6 168514	2010 847639	3 435999	–	–	–	–
Solution	0 416105	6 277385	2010 828263	3 501398	1510 132380	-475 550938	-36 057596	1842 039860
Vecteur initial	0 501445	5 795347	2010 778143	0 937666	–	–	–	–
Solution	0 416105	6 277468	2010 828272	0 359701	1510 137454	-475 553871	36 057533	1842 039860
Vecteur initial	0 186440	5 970107	2011 231866	-0 098513	–	–	–	–
Solution	0 170101	7 501234	2011 488906	-2 642171	3071 727267	-1860 469470	40 404655	1859 920060
Vecteur initial	0 263123	7 887407	2011 964151	1 754966	–	–	–	–
Solution	0 170101	7 501234	2011 488905	3 641016	3071 722105	-1860 464589	40 404645	1859 920060

Vecteur initial	0.264453	7.791482	2012.081755	4.760783	—	—	—	—
Solution	0.170101	7.501234	2011.488906	6.782607	3071.724688	-1860.467032	-40.404651	1859.920060

Tableau 4.6 : Ajustement aux prix de l'or (USD) — Données brutes

Les tableaux 4.5 et 4.6 semblent suggérer que l'ajustement aux prix de l'or est plus sensible au vecteur initial que ne l'est l'ajustement au Dow Jones précédant le krach d'octobre 1987. Néanmoins, les solutions obtenues sont cohérentes dans la mesure où toutes celles issues d'ajustements effectués au logarithme des données offrent des solutions inadmissibles alors que toutes celles issues d'ajustements aux données brutes annoncent l'imminence d'un krach.

Les graphes issus des meilleurs ajustements aux prix de l'or — l'un au logarithme des données, l'autre aux données brutes — ressemblent à ceux obtenus pour l'ajustement au krach d'octobre 1987 pour le Dow Jones tels que représentés dans la figure 4.4. Par contre, alors que les graphes du krach d'octobre semblent bien illustrer des lois de puissance décorées par des oscillations log-périodiques, ceux des prix de l'or semblent davantage représenter une droite décorée par des oscillations log-périodiques ; cette constatation est particulièrement évidente sur le graphe du logarithme des données, ci-dessous.

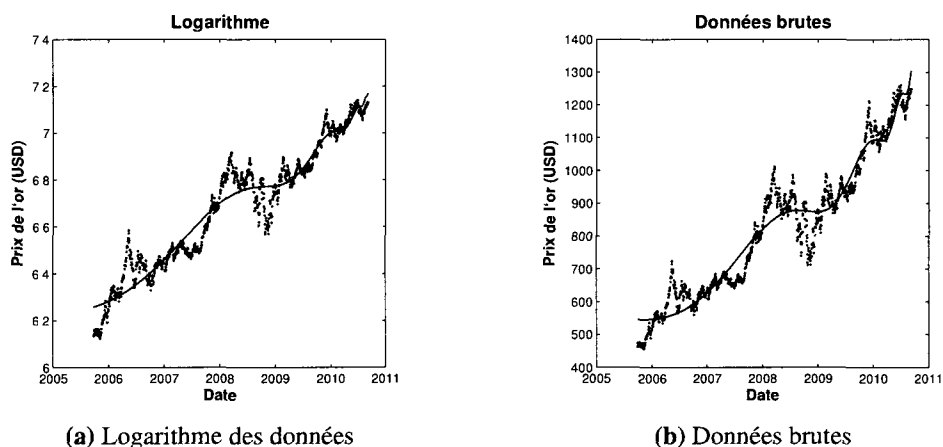


Figure 4.5 : Meilleurs ajustements aux prix de l'or

Il est donc très difficile de conclure sur l'imminence d'un krach au sein du marché de l'or. Néanmoins, l'ajustement aux données brutes avait prédit un krach au 30 octobre 2010 — ou au 28 juin 2011, si l'on se fie aux trois derniers ajustements ; il sera intéressant de voir ce qui arrivera sur ce marché au cours des prochains mois.

Discussion

Les résultats obtenus sont non concluants étant donnée la disparité présente dans la littérature. Davantage d'essais seraient nécessaires pour conclure sur l'applicabilité du modèle à des données empiriques. Il faudrait notamment effectuer des ajustements à davantage de séries financières, incluant : des krachs historiques, des données présentant des tendances sans toutefois afficher de krach et des données ne présentant aucune tendance notable.

Les résultats obtenus semblent néanmoins cohérents : en utilisant ou non le logarithme des données, les valeurs finales des paramètres issues de l'ajustement sont confinées à des intervalles relativement fixes et restreints. La date prévue du krach, le cas échéant, est étonnamment stable.

Quoi qu'il en soit, le processus de détermination des paramètres « optimaux » est actuellement beaucoup trop compliqué ou alors beaucoup trop flou. Il faudrait réviser la littérature disponible sur le sujet et uniformiser certains aspects du processus d'ajustement.

4.3 Anomalies et problématiques

Malheureusement, plusieurs anomalies et problématiques entourent le modèle LPPL. La première problématique surgit d'ailleurs très tôt dans l'élaboration du modèle : Sornette *et al.* utilisent un bon nombre d'hypothèses qu'ils ne justifient pas, qui ne semblent pas justifiées ou alors qui ne sont pas vérifiables. Les hypothèses non justifiées telles que l'équation (4.14)

sont problématiques dans la mesure où le fondement même du modèle doit être remis en question. Cette problématique survient assez fréquemment dans les travaux de Sornette *et al.* et les justifications que les auteurs en fournissent lors de communications privées sont incomplètes, douteuses ou cherchent tout simplement à esquiver la problématique.

4.3.1 Hypothèses non vérifiables

Brée semble être la seule personne, autre que les auteurs, à avoir effectué une analyse relativement complète du modèle et à avoir publié ses travaux. Il est un fervent dénonciateur des hypothèses douteuses, non justifiées ou non vérifiables émises par les auteurs originaux et il s'efforce de les mettre à l'épreuve lorsque possible.

Mais les hypothèses non justifiées qui ne sont pas vérifiables sont-elles vraiment problématiques ? Tout modèle applicable à des données empiriques se doit d'être réducteur, mais s'il fonctionne, ses hypothèses simplificatrices, à défaut d'être vérifiables, sont justifiées. Un réel problème survient, toutefois, lorsqu'une hypothèse cause une incohérence.

4.3.2 Données brutes ou logarithme ?

Le modèle LPPL repose sur l'élaboration d'une formule représentant le logarithme des prix d'un marché. Or, dans leurs travaux, Sornette *et al.* ajustent très souvent le modèle aux données brutes plutôt qu'à leur logarithme. Ils essayent de justifier cette pratique de la façon suivante.

Supposons que la chute de prix durant le krach soit proportionnelle à l'amplitude du prix au-dessus de la valeur fondamentale plutôt que d'être directement proportionnelle au prix comme le suggère l'équation (4.5). Cette hypothèse englobe la notion que le prix d'un actif tend à retrouver sa juste valeur après s'en être préalablement écarté. En notant par p_1 la valeur

fondamentale, l'équation (4.5) devient donc :

$$dp(t) = \kappa h(t)(p(t) - p_1) dt. \quad (4.31)$$

Sornette *et al.* mentionnent que $p(t)$ peut être interprété comme étant la valeur en excès de l'action par rapport à sa valeur fondamentale ; il n'est pas du tout clair [24][27] s'ils considèrent uniquement le cas où $p_1 = 0$ (pensons ici à un actif n'offrant aucun dividende) ou s'ils adoptent cette définition en général.

De plus, Johansen et Sornette [27] introduisent l'hypothèse additionnelle voulant que l'augmentation du prix depuis le début de la bulle soit nettement inférieure à la différence entre le prix au début de la bulle et la valeur fondamentale, c'est-à-dire :

$$p(t) - p(t_0) \ll p(t_0) - p_1 \quad (4.32)$$

où t_0 est le moment du début de la bulle. Cette hypothèse a une répercussion directe sur la définition qu'ils adoptent d'une « bulle ».

Il est impossible de mettre à l'épreuve ces hypothèses étant donné qu'il n'y a aucune manière de déterminer la valeur fondamentale d'un actif. Par ailleurs, nous pouvons en déduire que le prix ne peut faire davantage que doubler au cours de la bulle, sans quoi nous aurions $p(t) > 2p(t_0)$ et la valeur fondamentale devrait être négative. Puisque la valeur fondamentale d'un actif ne peut être négative, nous allons considérer qu'il n'est pas convenable d'ajuster le modèle LPPL aux données brutes si le prix double au cours de la bulle. Il n'est cependant pas clair si Sornette *et al.* utilisent une formulation alternative du modèle afin d'effectuer des ajustements aux données brutes ou s'ils utilisent malgré tout la même formule[10][27].

4.3.3 Martingale et l'hypothèse des anticipations rationnelles

Sous l'hypothèse des anticipations rationnelles, Sornette *et al.* proposent l'équation (4.2) comme modèle sous-jacent au prix d'une action. En cherchant à incorporer au modèle la propriété de martingale, nous obtenons l'équation (4.5), notamment $dp(t) = \kappa h(t)p(t) dt$.

Bien que la propriété de martingale soit difficilement vérifiable, il est néanmoins possible de questionner la validité des hypothèses avancées par les auteurs.

Analysons les composantes de l'équation (4.5) : κ est une valeur positive, $p(t)$ est le prix de l'action au temps t et donc une valeur positive, et $h(t)$ est le taux de risque, qui est supposé être une densité de probabilité. Ainsi, tous les facteurs à droite de l'égalité sont non-négatifs, d'où dp doit être non-négatif. Sous l'hypothèse de martingale et en supposant que le taux de risque est une densité de probabilité, nous en déduisons donc d'après le modèle LPPL que le prix d'une action doit toujours augmenter au cours d'une bulle. Or, en pratique, ceci n'est pas souvent le cas. Par conséquent, et du fait que le modèle soit dérivé de l'équation (4.5), nous devons rejeter l'hypothèse voulant que $h(t)$ soit une densité de probabilité.

4.3.4 Sept paramètres et le surajustement

Une critique commune concernant le modèle LPPL est qu'en présence d'autant de paramètres, il y a une forte possibilité de surajustement. Le problème ne consiste pas à savoir si le modèle peut être ajusté à n'importe lequel ensemble de données financières, mais plutôt de déterminer comment distinguer ceux qui doivent être précurseurs de krachs de tous les autres. Sornette a d'ailleurs consacré beaucoup d'effort à tenter de justifier statistiquement l'importance de la présence de chacun des paramètres dans le modèle[55][65]. Toutefois, même en ignorant le problème du surajustement, il reste à déterminer *comment* bien ajuster le modèle.

4.3.5 Critères d'identification d'un krach

À la question de savoir comment distinguer les ensembles de données financières précurseurs de krachs des ensembles inintéressants, la solution la plus naturelle est de considérer le vecteur des paramètres qui minimise (au sens des moindres carrés) la différence entre les valeurs prédites par le modèle et les données réelles, et d'établir des critères sur les résidus afin de déterminer si l'ajustement est convenable (et précurseur d'un krach) ou non. Or, ce n'est pas

l'approche pour laquelle ont opté Sornette *et al.* ; ils ont préféré établir des intervalles — basés sur l'ajustement du modèle à des krachs passés — dans lesquels doivent se trouver les paramètres afin que l'ajustement indique l'approche d'un krach.

Pourquoi procéder d'une telle manière ? Certains chercheurs, comme Brée, dénoncent l'approche choisie par Sornette *et al.* par le fait qu'aucune justification ne semble évidente pour préférer cette méthode à l'utilisation d'une mesure de qualité de l'ajustement et que les auteurs ne font pas l'effort d'en fournir. Il semble peut-être plus prudent d'utiliser une mesure de qualité de l'ajustement afin de vérifier la validité de l'ajustement (le carré moyen des résidus en étant un bon exemple), mais certaines situations portent à réflexion : imaginons le cas où, sur une période donnée, le marché affiche une tendance haussière suivant une loi de puissance comportant très peu de bruit. Dans ce cas, l'ajustement du modèle aux données serait excellent. Toutefois, nous aurions $C \approx 0$ et la partie responsable de l'oscillation log-périodique — soit la partie qui selon Sornette est annonciatrice de la présence d'une bulle et donc de l'imminence d'un krach — serait absente. Ainsi, même en présence d'une mesure de qualité de l'ajustement, des contraintes additionnelles sont nécessaires pour départir les données d'intérêt des autres données. Ironiquement, ces contraintes sont exactement de la forme de celles fournies par Sornette *et al.* telles sont présentées dans le tableau 4.1 ; par ailleurs, notons que la condition $C \neq 0$ n'y apparaît pas !

En optant pour une telle approche, nous évitons les problèmes rattachés à l'utilisation d'une unique mesure de qualité de l'ajustement, mais d'autres questions surgissent : si les conditions ont été établies en fonction de données historiques, pouvons-nous réellement être assurés que ces conditions discrimineront à l'avenir tous les ensembles de données qui ne sont pas d'intérêt et *seulement* ceux-là ? Pensons ici au krach d'octobre 2008 et au « *flash crash*⁴ » du 6 mai 2010, deux événements qui présentent des particularités qui ne sont pas observables dans divers krachs historiques.

⁴Cette appellation est utilisée pour souligner le fait que le krach fut quasi-instantané et très éphémère.

4.3.6 Sensibilité de l'ajustement

Sensibilité à la fenêtre temporelle

La littérature disponible [19] concernant le modèle LPPL, ainsi que nos essais, suggèrent que l'ajustement du modèle est particulièrement sensible à la fenêtre temporelle utilisée. En particulier, la date de début de la fenêtre sous étude affecte énormément les valeurs des paramètres finaux ; la date de fin, selon les travaux de Sornette *et al.*, n'affecte pas particulièrement l'issue de l'ajustement. Mais s'agit-il là des valeurs des paramètres finaux (et de la date prédite du krach) ou alors du verdict final (krach ou non) ? Sornette *et al.* suggèrent que la date de début soit sélectionnée comme étant le prix le plus bas depuis le dernier krach. Or, dans leurs articles, il est mentionné que la date de début fut parfois modifiée manuellement afin de convenir davantage à leur perception d'une bulle économique. Malgré cela, ils incorporent dans leur processus une étape exhaustive qui consiste à mettre à l'essai une quantité remarquable de fenêtres temporelles afin de déterminer le meilleur ajustement possible ; il est peut-être là question de surajustement. Il faudrait plutôt, à notre avis, standardiser le processus d'identification de la date de début (ou possiblement les dates) à utiliser pour la fenêtre temporelle.

Pour un modèle dont le verdict dépend uniquement de l'ajustement, c'est un réel problème !

Un point à garder en tête : lorsque plusieurs fenêtres temporelles sont mises à l'essai, la valeur SE n'est pas adéquate pour comparer les ajustements. Il faut alors pondérer la mesure d'ajustement à la longueur de la fenêtre temporelle utilisée afin de pouvoir comparer tous les ajustements entre eux. Nous obtenons ainsi les nouvelles mesures $MSE = \frac{1}{N}SE$ et $RMSE = \sqrt{MSE}$; cette première valeur est le carré moyen des résidus.

Sensibilité à l'échelle temporelle ?

À ma connaissance, aucune étude ne considère la question de l'échelle temporelle utilisée. Y aurait-il une différence à l'ajustement ou au verdict final si les données utilisées correspon-

daient aux prix intrajournaliers plutôt qu'aux valeurs quotidiennes de fermeture à la bourse ?

Faute d'avoir les données à notre disposition et des outils informatiques suffisamment puissants (étant donnée la quantité d'information), nous n'avons malheureusement pas pu enquêter sur le sujet.

4.3.7 Limitations pratiques

Le modèle LPPL offre un avantage très intéressant par rapport à d'autres modèles de prédiction de krachs : une date critique. Alors que d'autres modèles ne font qu'indiquer si un krach aura lieu ou non, le modèle LPPL offre une approximation du moment où cet événement devrait avoir lieu. Un problème, néanmoins : c'est une approximation déterministe, pas probabiliste ! En effet, le modèle LPPL fournit une manière d'obtenir une unique date à titre indicatif, mais aucune manière concrète de déterminer la probabilité actuelle d'un krach. Bien sûr, le modèle fournit une définition explicite du taux de risque $h(t)$, mais il n'y a pas de moyen facile de passer de l'équation (4.1) à l'équation (4.15), même après avoir déterminé les paramètres « optimaux ». De plus, la section 4.3.3 réfute l'hypothèse voulant que $h(t)$ soit une densité de probabilité.

Afin d'obtenir un modèle probabiliste, Sornette *et al.* ont récemment cherché [61] à établir des intervalles de confiance autour de la date critique t_c . Pour en arriver à leurs fins, ils ont décidé de faire appel à la méthode de *bootstrap par rééchantillonnage des résidus*. D'une série de données originales, ils génèrent dix (10) ensembles de données synthétiques. Chacune de ces nouvelles séries est soumise à un ajustement du modèle LPPL et la formule résultante est interpolée sur une période de six mois suivant la date de la dernière donnée disponible dans la série. De cette nouvelle famille de données, ils établissent des intervalles de confiance pour la date critique t_c .

Certains chercheurs doutent que les statistiques issues de cette approche soient réellement significatives[8]. Il existe un très gros problème philosophique concernant l'application que

Sornette *et al.* font de la méthode du bootstrap : comment peut-on ériger un intervalle de confiance autour de la valeur d'un paramètre issu de l'ajustement d'un modèle à un ensemble de données sans même savoir si ce modèle décrit adéquatement les données sous étude. En effet, les statistiques issues de cette approche ne peuvent être significatives qu'en admettant à priori la validité du modèle utilisé. La méthode de bootstrap par rééchantillonnage des résidus est utile pour analyser une propriété de la série originale qui ne repose pas sur la validité du modèle sous étude : la moyenne ou la variance de la distribution, à titre d'exemple. Bien sûr, toute probabilité suppose un modèle sous-jacent, mais les probabilités issues de l'application de la méthode du bootstrap au modèle LPPL reposent davantage sur la qualité des ajustements effectués que sur les propriétés des données originales.

Afin d'obtenir des probabilités réellement significatives concernant l'évolution d'un marché, il vaudrait mieux investir dans l'élaboration d'une théorie probabiliste cohérente avec le modèle LPPL, ou alors établir un nouveau modèle purement stochastique qui permettrait l'utilisation d'outils probabilistes.

4.4 Efforts et développements récents

Malgré toutes les lacunes que présente le modèle aux niveaux théorique et pratique, il a néanmoins déjà su démontrer une certaine efficacité dans l'identification de krachs passés ; même Brée admet que l'ajustement du modèle à certaines données est surprenant. Mais qu'en est-il lorsqu'il s'agit de krachs qui ne sont pas encore survenus ?

Afin de démontrer l'efficacité réelle du modèle, Sornette, Woodard, Fedorovsky, Reimann, Woodard et Zhou [61] ont entrepris ce qu'ils appellent la « *Financial Bubble Experiment* »⁵. Ils ont sélectionné quatre actions pour lesquelles ils anticipaient un changement de régime au cours des prochains six mois, ces prédictions étant issues de plusieurs méthodes qu'ils refusent de divulguer dont des ajustements du modèle LPPL aux données disponibles à ce

⁵Une traduction possible est « l'expérience de bulle économique ».

moment. Malgré tout, leur crédibilité est en jeu : ils publient « au fur et à mesure »⁶ les résultats de leurs analyses et leurs anticipations du moment.

Apparemment, les quatre actions sélectionnées ont effectivement présenté un changement de régime, mais l'issue de l'expérience ne sera pas discutée ici : il suffit de consulter le site Web de la « Financial Crisis Observatory » à la *Eidgenössische Technische Hochschule Zürich* pour plus de détails. Profitons toutefois de l'occasion pour applaudir leur audace et leur courage.

4.5 Discussion

Bien que le modèle LPPL semble offrir une voie intéressante en pratique, son fondement théorique est souvent basé sur des hypothèses dont la validité ne peut être vérifiée. Les notions utilisées dans l'application du modèle sont également mal définies : pensons ici à l'identification de la fenêtre temporelle sur laquelle se produit une bulle économique. Les publications de Sornette *et al.*, trop souvent dépourvues de certaines informations clés, contribuent à brouiller davantage les étapes déjà obscures dans leur application du modèle : pensons notamment à l'utilisation des données brutes plutôt que leur logarithme alors que le fondement théorique du modèle repose sur l'utilisation du logarithme. L'avancement et l'applicabilité d'un modèle tendent à prospérer davantage lorsqu'il y a partage d'information et suivant l'effort de plusieurs individus ou groupes d'individus. Or, l'information contenue dans les publications de Sornette est généralement restreinte et incomplète : il est souvent très difficile de déterminer si les méthodes utilisées par Sornette sont valables et pourquoi. Ceci rend la tâche d'autant plus fastidieuse pour les gens qui tentent de reproduire et de valider ses résultats de façon indépendante. Brée, qui est probablement l'un des critiques les plus durs des travaux de Sornette, a d'ailleurs eu besoin d'entrer en communication avec Johansen à plusieurs reprises afin d'en arriver à son analyse du modèle LPPL.

Suite à l'annonce de l'expérience effectuée par Sornette *et al.* afin de crédibiliser le modèle

⁶Il s'agit en fait de documents qu'ils ont publiés et encryptés au moment de leurs prédictions et qu'ils ont rendus publics après plusieurs mois.

LPPL, il fut demandé à Sornette s'il avait l'intention de déposer une demande de brevet pour sa méthode comme il l'avait fait au cours des années 1990 pour un modèle de prédiction de fissures dans les matériaux de fusées. Sa réponse fut la suivante :

Non, je n'ai pas l'intention de faire cela cette fois-ci. Si vous désirez breveter quelque chose, il vous faut divulguer vos méthodes, mais nous gardons une partie du savoir-faire technique pour nous-mêmes. Pour nous, c'est une mesure de notre succès à savoir combien d'institutions financières et d'experts sont entrés en communication avec nous dans le but de travailler avec nous. Ils stimulent également notre recherche avec leurs questions. [57]

Chapeau à Mandelbrot qui croyait que le savoir appartient à toutes et à tous, et qui cherchait la contribution de quiconque voulait bien s'y attarder pour faire avancer sa théorie.

CHAPITRE 5

Modélisation stochastique

Le problème le plus important concernant le modèle LPPL a été soulevé dans le précédent chapitre : alors que l'identification d'un krach repose sur l'estimation des paramètres de l'ajustement du modèle aux données empiriques, celle-ci s'avère beaucoup trop difficile et imprécise. À cet effet, ma suggestion était d'établir une théorie probabiliste cohérente avec le modèle LPPL, ou alors de développer un nouveau modèle purement stochastique afin que nous puissions utiliser les outils du calcul stochastique et du calcul d'Itô. Or, cette seconde option est exactement ce que Didier Sornette et Li Lin ont fait : le 10 novembre 2009, un document intitulé « Diagnostics of Rational Expectation Financial Bubbles with Stochastic Mean-Reverting Termination Times » [33] est apparu sur *arXiv*. Dans cet article, les auteurs proposent deux nouveaux modèles basés sur la théorie des anticipations rationnelles dont les dynamiques globales sont représentées par des processus purement stochastiques. Ces modèles préservent certaines caractéristiques intrinsèques au modèle LPPL : la présence de renforcement positif influençant la dynamique du prix et l'existence d'un temps critique \tilde{t}_c . Cette approche a l'avantage d'ignorer le fonctionnement interne du système.

Les deux nouveaux modèles présentés par Lin et Sornette cherchent à surmonter les problématiques que sont l'estimation des paramètres dans l'ajustement du modèle LPPL et le caractère purement déterministe de ce dernier. Toutefois, les calculs présents dans l'article

sont erronés : nous apportons donc sommairement les corrections nécessaires et nous présentons les calculs rectifiés.

5.1 Premier modèle

En période de bulle économique, nous supposons que la dynamique du prix est décrite par

$$dp = \mu p^m (1 + \delta(p, t)) dt + \sigma p^m dW \quad (5.1)$$

où W est un processus de Wiener. Lin et Sornette montrent que, lorsque $m > 1$, le prix diverge en temps fini, plus précisément au temps critique \tilde{t}_c . Dans l'équation (5.1), μ est le taux de rendement instantané, σ est la volatilité des rendements et $\delta(p, t)$ est un processus à déterminer qui gouverne le comportement de \tilde{t}_c .

Lorsque $m = 1$ et $\delta(p, t) = 0$, l'équation (5.1) devient

$$dp = \mu p dt + \sigma p dW \quad (5.2)$$

et nous retrouvons le modèle de Black-Scholes tel que défini par l'équation (1.17).

Voilà une propriété du modèle qui est particulièrement attrayante : l'équation (5.2) dit que le modèle de Black-Scholes, notamment le modèle adéquatement utilisé de nos jours lors de périodes dépourvues de bulles économiques, n'est qu'un cas particulier d'un modèle plus général pouvant dicter la dynamique du marché à tout moment, en présence comme en l'absence de bulles. C'est une théorie élégamment conciliante et purement stochastique !

5.1.1 Cas simplifié

Considérons le cas particulier où $\delta(p, t) = \sigma = 0$. L'équation (5.1) devient alors

$$dp = \mu p^m dt. \quad (5.3)$$

Cherchons la solution $p(t)$ à cette dernière équation.

De l'équation (5.3), nous déduisons

$$p^{-m} dp = \mu dt.$$

En intégrant de $t_0 = 0$ à t , nous obtenons

$$\left[\frac{-1}{m-1} p(t')^{-(m-1)} \right]_{t'=0}^{t'=t} = \int_0^t p(t')^{-m} dp dt' = \int_0^t \mu dt' = [\mu t']_{t'=0}^{t'=t},$$

ou encore

$$\frac{-1}{m-1} p^{-(m-1)} + \frac{1}{m-1} p(0)^{-(m-1)} = \mu t.$$

Posons $\beta = \frac{1}{m-1}$ et notons $p_0 := p(0)$. Nous avons

$$-\beta p^{-(m-1)} = -\beta p_0^{-(m-1)} + \mu t,$$

d'où

$$p^{-(m-1)} = \frac{\mu}{\beta} \left(\frac{\beta}{\mu} p_0^{-(m-1)} - t \right)$$

et finalement

$$p(t) = \left(\frac{\beta}{\mu} \right)^\beta \left(\frac{p_0^{-(m-1)}}{\mu(m-1)} - t \right)^{-\beta}.$$

En posant $K := \left(\frac{\beta}{\mu} \right)^\beta$ et $t_c := \frac{p_0^{-(m-1)}}{\mu(m-1)}$, nous concluons que la solution à l'équation (5.3) est donnée par

$$p(t) = K(t_c - t)^{-\beta}. \quad (5.4)$$

5.1.2 Cas général

Motivés par la forme de la solution du cas simplifié, nous cherchons à déterminer $\delta(p, t)$ de sorte que la solution à l'équation (5.1) soit de la forme

$$p(t) = K(\tilde{T}_c - t)^{-\beta}.$$

La définition que nous fournissons de $\delta(p, t)$ diffère de celle fournie par Lin et Sornette afin de corriger la prochaine proposition. Nous postulons la forme suivante pour le processus gouvernant $\delta(p, t)$:

$$\delta(p, t) = \alpha \tilde{t}_c(t) + \frac{1}{2} m(\sigma^2/\mu) [p(t)]^{m-1} \quad (5.5)$$

où

$$d\tilde{t}_c(t) = -\alpha \tilde{t}_c(t) dt - (\sigma/\mu) dW \quad (5.6)$$

est un processus d'Ornstein-Uhlenbeck de moyenne zéro. Le processus de Wiener dans l'équation (5.6) est le même que celui dans l'équation (5.1).

Proposition 5.1.1

Avec la dynamique de $\delta(p, t)$ décrite par les équations (5.5) et (5.6), la solution à l'équation (5.1) peut être écrite sous la forme

$$p(t) = K(\tilde{T}_c - t)^{-\beta} \quad (5.7)$$

où

$$\beta = \frac{1}{m-1}, \quad K = \left(\frac{\beta}{\mu}\right)^\beta, \quad T_c = (\beta/\mu)p_0^{-1/\beta}, \quad \tilde{T}_c = T_c + \tilde{t}_c. \quad (5.8)$$

Démonstration. Nous devons avoir recours ici au calcul d'Itô.

Posons $n := m - 1$ et $z := p^{-n}$. Plus explicitement, $z(t, p(t)) = p(t)^{-n}$. Le lemme 1.4.8 (lemme d'Itô) et l'équation (5.1) nous donnent

$$dz(t, p) = z_t(t, p) dt + z_p(t, p) dp + \frac{1}{2} z_{pp}(t, p) (\sigma p^m)^2 dt. \quad (5.9)$$

Or, $z_t(t, p) = \frac{\partial}{\partial t} p^{-n} = 0$, d'où

$$dz = z_p dp + \frac{1}{2} z_{pp} (\sigma p^m)^2 dt. \quad (5.10)$$

De plus,

$$\begin{aligned} z_p &= -n p^{-(n+1)} = -n p^{-m}, \\ z_{pp} &= n(n+1) p^{-(n+2)} = n(n+1) p^{-(m+1)}, \end{aligned}$$

d'où

$$dz = -np^{-m}dp + \frac{1}{2}n(n+1)p^{-(m+1)}\sigma^2 p^{2m} dt. \quad (5.11)$$

En substituant l'équation (5.1) dans cette dernière, nous obtenons

$$\begin{aligned} dz &= -np^{-m}[\mu p^m(1 + \delta(p, t))dt + \sigma p^m dW] + \frac{1}{2}n(n+1)p^{-(m+1)}\sigma^2 p^{2m} dt \\ &= -n\mu(1 + \delta(p, t))dt + n\sigma dW + \frac{1}{2}n(n+1)\sigma^2 p^{m-1} dt \\ &= -n\mu \left[1 + \delta(p, t) - \frac{1}{2}m \frac{\sigma^2}{\mu} p^{m-1} \right] dt - n\sigma dW. \end{aligned}$$

L'équation (5.5) nous donne alors

$$\begin{aligned} dz &= -n\mu \left[1 + \left(\alpha \tilde{t}_c + \frac{1}{2}m \frac{\sigma^2}{\mu} p^{m-1} \right) - \frac{1}{2}m \frac{\sigma^2}{\mu} p^{m-1} \right] dt - n\sigma dW \\ &= -n\mu[1 + \alpha \tilde{t}_c] dt - n\sigma dW \\ &= -n\mu dt + n\mu \left(-\alpha \tilde{t}_c dt - \frac{\sigma}{\mu} dW \right). \end{aligned}$$

En remplaçant la valeur de $d\tilde{t}_c$ donnée par l'équation (5.6), nous trouvons

$$\begin{aligned} d(p^{-n}) &= dz = -n\mu dt + n\mu d\tilde{t}_c \\ &= n\mu(d\tilde{t}_c - dt). \end{aligned}$$

En intégrant de $t_0 = 0$ à t de chaque côté et en posant $p_0 := p(0)$, nous obtenons

$$p^{-n} - p_0^{-n} = n\mu(\tilde{t}_c(t) - t - \tilde{t}_c(0)).$$

Puisque \tilde{t}_c est distribué selon un processus d'Ornstein-Uhlenbeck de moyenne zéro, nous posons $\tilde{t}_c(0) := 0$, d'où

$$p^{-n} = n\mu \left(\tilde{t}_c - t + \frac{p_0^{-n}}{n\mu} \right)$$

et donc

$$p(t) = (n\mu)^{-1/n} \left(\tilde{t}_c - t + \frac{p_0^{-n}}{n\mu} \right)^{-1/n}.$$

En substituant $\beta = \frac{1}{m-1}$, $T_c = \frac{\beta}{\mu} p_0^{-1/\beta}$, $\tilde{T}_c = T_c + \tilde{t}_c$, $K = (\beta/\mu)^{-\beta}$ et $n = m - 1$ dans cette dernière équation, nous retrouvons l'équation (5.7). \square

5.2 Second modèle

Posons $y(t) := \ln p(t)$ où $p(t)$ est le prix d'une action en temps de bulle économique et supposons que la dynamique de $y(t)$ est décrite par les processus

$$dy = x(1 + \gamma(x, t)) dt + (\sigma/\mu)x dW \quad (5.12)$$

$$dx = \mu x^m(1 + \delta(x, t)) dt + \sigma x^m dW \quad (5.13)$$

où $\gamma(x, t)$ et $\delta(x, t)$ sont des processus à déterminer.

5.2.1 Cas simplifié

Considérons le cas particulier où $\gamma(x, t) = \delta(x, t) = \sigma = 0$. Les équations (5.12) et (5.13) deviennent alors

$$dy = x dt \quad (5.14)$$

$$dx = \mu x^m dt. \quad (5.15)$$

L'équation (5.14) nous donne $\frac{dy}{dt} = x$, d'où

$$\frac{d^2y}{dt^2} = \frac{dx}{dt} = \mu x^m, \quad (5.16)$$

la dernière égalité découlant de l'équation (5.15). Ainsi, en substituant la valeur de x fournie par l'équation (5.14), nous obtenons

$$\frac{d^2y}{dt^2} = \mu \left(\frac{dy}{dt} \right)^m. \quad (5.17)$$

En allégeant la notation, cette dernière équation se réécrit sous la forme

$$y'' = \mu(y')^m \quad (5.18)$$

Cherchons la solution $y(t)$ à cette dernière équation. De l'équation (5.18), nous déduisons

$$(y')^{-m} y'' = \mu.$$

En intégrant des deux côtés, nous obtenons

$$\frac{-1}{m-1} (y')^{-(m-1)} = \mu t + c_1$$

où c_1 est une constante réelle. Posons $\beta = \frac{1}{m-1}$. Ainsi,

$$-\beta (y')^{-1/\beta} = \mu t + c_1, \quad (5.19)$$

d'où

$$y' = (1/\beta)^{-\beta} (-c_1 - \mu t)^{-\beta}. \quad (5.20)$$

En intégrant à nouveau des deux côtés, nous obtenons

$$\begin{aligned} y &= \left(\frac{1}{\beta}\right)^{-\beta} \left(\frac{-1}{\mu}\right) \left(\frac{-1}{\beta-1}\right) (-c_1 - \mu t)^{-(\beta-1)} + c_2 \\ &= \left(\frac{-1}{1-\beta}\right) \left(\frac{1}{\beta}\right)^{-\beta} \left(\frac{1}{\mu}\right) (-c_1 - \mu t)^{1-\beta} + c_2 \\ &= \left(\frac{-1}{1-\beta}\right) \left(\frac{1}{\beta}\right)^{-\beta} \left(\frac{1}{\mu}\right) \mu^{1-\beta} \left(\frac{-c_1}{\mu} - t\right)^{1-\beta} + c_2 \\ &= \left(\frac{-1}{1-\beta}\right) \left(\frac{1}{\beta}\right)^{-\beta} \mu^{-\beta} \left(\frac{-c_1}{\mu} - t\right)^{1-\beta} + c_2 \\ &= \left(\frac{-1}{1-\beta}\right) \left(\frac{\mu}{\beta}\right)^{-\beta} \left(\frac{-c_1}{\mu} - t\right)^{1-\beta} + c_2. \end{aligned} \quad (5.21)$$

Déterminons maintenant les valeurs des constantes c_1 et c_2 .

L'équation (5.19) nous donne

$$(\beta/\mu) (y')^{-1/\beta} + t = \frac{-c_1}{\mu}.$$

En évaluant cette dernière équation en $t = 0$, nous obtenons

$$\frac{-c_1}{\mu} = (\beta/\mu) \left(\frac{dy}{dt} \Big|_{t=0} \right)^{-1/\beta} = T_c \quad (5.22)$$

Posons $B = \frac{1}{1-\beta} (\mu/\beta)^{-\beta}$. Ainsi, l'équation (5.21) devient

$$y = c_2 - B(T_c - t)^{1-\beta}.$$

En évaluant cette dernière équation en $t = T_c$, nous obtenons

$$y(T_c) = c_2$$

et donc, en posant $A = y(T_c)$, nous aboutissons à la solution désirée, notamment

$$y = A - B(T_c - t)^{1-\beta}. \quad (5.23)$$

5.2.2 Cas général

Motivés par la forme de la solution du cas simplifié, nous cherchons à déterminer $\gamma(x, t)$ et $\delta(x, t)$ de sorte que la solution aux équations (5.12) et (5.13) soit de la forme

$$y(t) = A - B(T_c + \tilde{t}_c(t) - t)^{1-\beta}.$$

Les définitions que nous fournissons de $\gamma(x, t)$ et $\delta(x, t)$ diffèrent de celles fournies par Lin et Sornette afin de corriger la prochaine proposition. Nous postulons les formes suivantes pour les processus gouvernant $\gamma(x, t)$ et $\delta(x, t)$:

$$\gamma(x, t) = \alpha \tilde{t}_c(t) + \frac{1}{2} (\sigma^2/\mu) [x(t)]^{m-1} \quad (5.24)$$

$$\delta(x, t) = \alpha \tilde{t}_c(t) + m \frac{1}{2} (\sigma^2/\mu) [x(t)]^{m-1} \quad (5.25)$$

où

$$d\tilde{t}_c(t) = -\alpha \tilde{t}_c(t) dt - (\sigma/\mu) dW \quad (5.26)$$

est un processus d'Ornstein-Uhlenbeck de moyenne zéro. Le processus de Wiener dans l'équation (5.26) est le même que celui des équations (5.12) et (5.13).

Remarque 5.2.1 - En considérant uniquement les équations (5.13), (5.25) et (5.26), nous retrouvons précisément le premier modèle.

Proposition 5.2.2

Avec les dynamiques de $\gamma(x, t)$ et $\delta(x, t)$ décrites par les équations (5.24) et (5.25) respectivement ainsi que l'équation (5.26), la solution aux équations (5.12) et (5.13) peut être écrite

sous la forme

$$y(t) = A - B(T_c + \tilde{t}_c - t)^{1-\beta} \quad (5.27)$$

où

$$\beta = \frac{1}{m-1}, \quad A = y(T_c), \quad B = \frac{1}{1-\beta}(\beta/\mu)^\beta, \quad T_c = (\beta/\mu)x_0^{1/\beta}, \quad x_0 = x(0). \quad (5.28)$$

Démonstration. Nous vérifions ici que l'équation (5.27) est bien la solution au système d'équations stochastiques fournies.

En appliquant le lemme (1.4.8) (lemme d'Itô) à l'équation (5.27), nous obtenons

$$dy(t, \tilde{t}_c) = y_t(t, \tilde{t}_c) dt + y_{\tilde{t}_c}(t, \tilde{t}_c) d\tilde{t}_c + \frac{1}{2} y_{\tilde{t}_c \tilde{t}_c}(t, \tilde{t}_c) (-\sigma/\mu)^2 dt. \quad (5.29)$$

En dérivant l'équation (5.27) par rapport aux variables appropriées, nous obtenons

$$\begin{aligned} y_t &= (1-\beta)B(T_c + \tilde{t}_c - t)^{-\beta} \\ y_{\tilde{t}_c} &= -(1-\beta)B(T_c + \tilde{t}_c - t)^{-\beta} \\ y_{\tilde{t}_c \tilde{t}_c} &= \beta(1-\beta)B(T_c + \tilde{t}_c - t)^{-\beta-1} \end{aligned}$$

d'où

$$\begin{aligned} dy &= (1-\beta)B(T_c + \tilde{t}_c - t)^{-\beta} dt - (1-\beta)B(T_c + \tilde{t}_c - t)^{-\beta} d\tilde{t}_c \\ &\quad + \frac{1}{2}\beta(1-\beta)B(T_c + \tilde{t}_c - t)^{-\beta-1}(\sigma/\mu)^2 dt. \end{aligned}$$

En substituant $B = \frac{1}{1-\beta}(\beta/\mu)^\beta$, nous avons

$$\begin{aligned} dy &= (\beta/\mu)^\beta (T_c + \tilde{t}_c - t)^{-\beta} dt - (\beta/\mu)^\beta (T_c + \tilde{t}_c - t)^{-\beta} d\tilde{t}_c \\ &\quad + \frac{1}{2}\beta(\beta/\mu)^\beta (T_c + \tilde{t}_c - t)^{-\beta-1}(\sigma/\mu)^2 dt. \end{aligned}$$

Posons

$$Z(t) = (\beta/\mu)^\beta (T_c + \tilde{t}_c - t)^{-\beta}. \quad (5.30)$$

Ainsi,

$$dy = Z dt - Z d\tilde{t}_c + \frac{1}{2}\beta Z (T_c + \tilde{t}_c - t)^{-1} (\sigma/\mu)^2 dt,$$

mais

$$\begin{aligned}\mu Z^{1/\beta} &= \mu \left[(\beta/\mu)^\beta (T_c + \tilde{t}_c - t)^{-\beta} \right]^{1/\beta} \\ &= \mu(\beta/\mu)(T_c + \tilde{t}_c - t)^{-1} \\ &= \beta(T_c + \tilde{t}_c - t)^{-1}\end{aligned}$$

de telle sorte que

$$\begin{aligned}dy &= Z dt - Z d\tilde{t}_c + \frac{1}{2} Z Z^{1/\beta} \mu (\sigma/\mu)^2 dt \\ &= Z [dt - d\tilde{t}_c + \frac{1}{2} (\sigma^2/\mu) Z^{1/\beta} dt].\end{aligned}$$

En substituant la valeur de $d\tilde{t}_c$ telle que définie par l'équation (5.26), nous obtenons

$$\begin{aligned}dy &= Z [dt + \alpha \tilde{t}_c dt + (\sigma/\mu) dW + \frac{1}{2} (\sigma^2/\mu) Z^{1/\beta} dt] \\ &= Z [1 + \alpha \tilde{t}_c + \frac{1}{2} (\sigma^2/\mu) Z^{m-1}] dt + Z (\sigma/\mu) dW.\end{aligned}$$

De la proposition 5.1.1, il est clair que $Z(t)$ défini par l'équation (5.30) est solution à l'équation (5.13) lorsque $\delta(x, t)$ et \tilde{t}_c sont définis par les équations (5.25) et (5.26) respectivement.

Ainsi, $Z = x$.

Nous avons

$$\begin{aligned}dy &= x \left(1 + \alpha \tilde{t}_c + \frac{1}{2} \frac{\sigma^2}{\mu} x^{m-1} \right) dt + (\sigma/\mu) x dW \\ &= x(1 + \gamma(x, t)) dt + (\sigma/\mu) x dW\end{aligned}$$

et donc $y(t)$ défini par les équations (5.27) et (5.28) est la solution désirée. \square

CONCLUSION

Avec l'émergence de branches telles l'éconophysique et la finance comportementale au cours des dernières décennies, de nouvelles méthodes d'analyse et de modélisation de marchés boursiers ont vu le jour. Bien qu'elles offrent toutes une perspective intéressante sur le sujet, aucune de ces approches n'a encore abouti à une théorie universelle et cohérente.

La théorie actuelle la plus prometteuse, si l'on se fie à ses auteurs, est certes celle que nous avons nous-mêmes mis sous étude au chapitre 4 : le modèle LPPL. Introduit en 1996, ce modèle est vanté depuis plusieurs années comme étant une méthode très efficace pour identifier l'arrivée prochaine d'un krach et pour en fournir une date approximative. Nos essais, toutefois, sont loin d'être aussi concluants. Nous avons décelé plusieurs problèmes et anomalies avec l'application de ce modèle à des données réelles, et toute conclusion au sujet des tests que nous avons effectués s'avère impossible étant donné l'incomplétude et l'incohérence parfois présente dans la littérature. Bien qu'il soit impossible d'offrir un verdict entièrement objectif sur l'applicabilité du modèle, son étude ne s'avère pas pour autant inutile. En effet, la majorité des gens mentionnés dans le mémoire ont reçu un « prix Nobel de l'économie », malgré leurs travaux qui sont maintenant fortement critiqués. Des théories qui s'annoncent prometteuses aujourd'hui pourraient être la risée de la prochaine décennie, et inversement. Quoi qu'il en soit, il demeure primordial d'émettre des hypothèses et des idées qui influencent le monde, qui stimulent la discussion et les débats, qu'elles soient erronées ou non : la critique qui en découle et les constatations empiriques qui en résultent contribuent grandement à l'avancement de notre connaissance collective.

Bien que nous soyons d'avis que le modèle LPPL mérite qu'on s'y attarde plus longuement (et certainement qu'on prenne le temps de rectifier la littérature sur le sujet), nous croyons également que la modélisation stochastique offre une alternative particulièrement intéressante, ne serait-ce que par sa nature. Si nous ne pouvons affirmer que cette dernière approche offrira une solution ultime à un problème aussi important, nous avons néanmoins la certitude qu'elle a beaucoup à nous apprendre sur un sujet que nous connaissons trop peu : la dynamique des systèmes dont le fonctionnement interne est gouverné par des règles floues et sujettes aux changements. En particulier, un modèle purement stochastique saura-t-il englober adéquatement la notion de « tendance » issue d'un comportement social collectif ?

L'étude d'un marché boursier introduit une difficulté qui n'apparaît pas dans plusieurs autres domaines : même si un modèle nous offre une connaissance de l'issue du marché, rien ne nous garantit la compréhension du mécanisme interne. En effet, le fonctionnement interne est si complexe que l'information qui nous est fournie par l'issue des événements est probablement insuffisante pour en conclure quoi que ce soit. À cet effet, bien qu'il puisse s'avérer impossible de modéliser fidèlement le mécanisme interne d'un marché boursier, nous avons la conviction qu'il est possible d'en modéliser l'issue, ne serait-ce que d'un point de vue probabiliste. Voilà pourquoi nous proposons la modélisation stochastique.

Or, un problème encore plus fondamental s'impose : l'issue d'un marché est une conséquence directe des actions des individus qui le composent. Même en présence d'un outil qui nous permettrait de déceler un éventuel effondrement de l'économie, il demeure entre les mains des constituants du marché de décider ce qui adviendra réellement...

Dans un monde où l'excès est recherché bien avant l'équilibre, seul le temps pourra nous révéler ce que l'être humain saura faire de ce système qu'il a créé.

Bibliographie

- [1] ABKEN, P., AND NANDI, S. Options and volatility. *Economic Review*, Dec (1996), 21–35.
- [2] ANDREASSEN, P., AND KRAUS, S. Judgmental prediction by extrapolation. *Harvard University Mimeo* (1988).
- [3] BACHELIER, L. Théorie de la spéculation. *Annales Scientifiques de L'École Normale Supérieure* 17 (1900), 21–88. Réimprimé dans P. H. Cootner (ed), 1964, *The Random Character of Stock Market Prices*, Cambridge, Massachusetts MIT Press.
- [4] BASU, S. Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios : A Test of the Efficient Market Hypothesis. *Journal of Finance* 32, 3 (1977), 663–682.
- [5] BLACK, F., AND SCHOLES, M. The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy* 81, 3 (1973), 637–654.
- [6] BOLLERSLEV, T. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31, 3 (1986), 307–327.
- [7] BRÉE, DAVID S., C. D., AND PEIRANO, P. Prediction accuracy and sloppiness of log-periodic functions, 2010, arXiv :q-fin.GN/1006.2010v1.
- [8] BRÉE, D., CHALLET, D., AND PEIRANO, P. Prediction accuracy and sloppiness of log-periodic functions, 2010, arXiv :q-fin.ST/1006.2010v1.
- [9] BRÉE, D., AND JOSEPH, N. Log Periodic Power Law fits to financial crashes : a preliminary replication. *Econophysics Colloquium and beyond* (2007), 27–29.

- [10] BRÉE, D., AND JOSEPH, N. Fitting the Log Periodic Power Law to Financial crashes : a critical analysis, 2010, arXiv :q-fin.ST/1002.1010v1.
- [11] CAMPBELL, J., LO, A., MACKINLAY, A., AND LO, A. *The Econometrics of Financial Markets*. Princeton University Press, 1996.
- [12] CHANG, G., AND FEIGENBAUM, J. A Bayesian analysis of log-periodic precursors to Financial crashes. *Quantitative Finance* 6 (2006), 15–36.
- [13] COOTNER, P. *The random character of stock market prices*, rev. ed ed. M.I.T. Press, Cambridge Massachusetts, 1964.
- [14] DERRIDA, B., DEŠEZE, L., AND ITZYKSON, C. Fractal structure of zeros in hierarchical models. *Journal of Statistical Physics* 33 (1983), 559.
- [15] EINSTEIN, A. Über die von der molekularkinetischen Theorie der Wärme gefordert Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik* 17 (1905), 549–560.
- [16] ENGLE, R. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* 50, 4 (1982), 987–1007.
- [17] FAMA, E. The Behavior of Stock-Market Prices. *The Journal of Business* 38, 1 (1965), 34–105.
- [18] FAMA, E. Efficient Capital Markets : A Review of Theory and Empirical Work. *The Journal of Finance* 25, 2 (1970), 383–417.
- [19] FEIGENBAUM, J. A Statistical Analysis of Log-Periodic Precursors to Financial Crashes, 2001, arXiv :cond-mat/0101031.
- [20] GLOVER, F., AND LAGUNA, M. *Tabu Search*, 1 ed. Springer, June 1998.
- [21] HENTSCHEL, L. All in the family nesting symmetric and asymmetric garch models. *Journal of Financial Economics* 39, 1 (1995), 71–104.
- [22] INGEN-HOUSZ, J. Vermischte Schriften physisch-medizinischen Inhaltes, 1784.
- [23] JOHANSEN, A. Comment on Recent Claims by Sornette and Zhou. *SSRN eLibrary* (2003).

- [24] JOHANSEN, A., LEDOIT, O., AND SORNETTE, D. Crashes as critical points. *International Journal of Theoretical and Applied Finance* 3 (2000), 219–255.
- [25] JOHANSEN, A., AND SORNETTE, D. Critical Crashes, 1999, arXiv :cond-mat/9901035v1.
- [26] JOHANSEN, A., AND SORNETTE, D. The Nasdaq crash of April 2000 : Yet another example of log-periodicity in a speculative bubble ending in a crash. *The European Physical Journal B* 17 (2000), 319–328, arXiv :cond-mat/0004263.
- [27] JOHANSEN, A., AND SORNETTE, D. Bubbles and anti-bubbles in Latin-American, Asian and western stock markets : An empirical study. *International Journal of Theoretical and Applied Finance* 4 (2001), 853–920.
- [28] JOHANSEN, A., SORNETTE, D., AND LEDOIT, O. Predicting Financial Crashes using discrete scale invariance. *Journal of Risk* 1 4 (1999), 5–32.
- [29] KELLY, C. Meet the man whose big idea felled Wall Street. thestar.com, Mars 2009. Récupéré le 2010-09-17 à l’adresse <http://www.thestar.com/article/604033>.
- [30] KOTZ, S., SALINETTI, G., AND DALL’AGLIO, G. *Advances in probability distributions with given marginals : beyond the copulas*. Kluwer Academic Publishers, 1991.
- [31] LI, D. X. On Default Correlation : A Copula Function Approach. *Journal of Fixed Income* 9, 4 (2000), 43–54.
- [32] LIN, L., AND SORNETTE, D. Diagnostics of Rational Expectation Financial Bubbles with Stochastic Mean-Reverting Termination Times, 2009, arXiv :q-fin.GN/0911.1921.
- [33] LIN, L., AND SORNETTE, D. Diagnostics of Rational Expectation Financial Bubbles with Stochastic Mean-Reverting Termination Times. arXiv :q-fin.GN/0911.1921.
- [34] MANDELBROT, B. The Variation of Certain Speculative Prices. *The Journal of Business* 36, 4 (1963), 394–419.
- [35] MANDELBROT, B. Une classe de processus stochastiques homothétiques à soi. Application à la loi climatologique de H. E. Hurst. *Comptes Rendus (Paris)* 260 (1965), 3274–3277.

- [36] MANDELBROT, B. Possible Refinements of the Lognormal Hypothesis Concerning the Distribution of Energy dissipation in Intermittent Turbulence. *Statistical Models and Turbulence* (1972), 333.
- [37] MANDELBROT, B. *Fractals : Form, Chance and Dimension*. W. H. Freeman & Company, San Francisco, 1977.
- [38] MANDELBROT, B. *The Fractal Geometry of Nature*. W. H. Freeman, 1982.
- [39] MANDELBROT, B. *Fractals and scaling in finance*. Springer, 1997.
- [40] MANDELBROT, B. *Gaussian Self-Affinity and Fractals — Globality, the Earth, 1/f Noise, and R/S*. Springer, 2002.
- [41] MANDELBROT, B., FISHER, A., AND CALVET, L. E. A Multifractal Model of Asset Returns. Cowles Foundation Discussion Papers 1164, Cowles Foundation, Yale University, 1997.
- [42] MANDELBROT, B., AND HUDSON, R. *The (mis)Behavior of Markets*. Basic Books, New York, 2004.
- [43] MANDELBROT, B., AND VAN NESS, J. Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review* 10, 4 (1968), 422–437.
- [44] MERTON, R. Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science* 4, 1 (1973), 141–183.
- [45] MUKHERJEE, T., AND HENDERSON, G. The Capital Budgeting Process : Theory and Practice. *Interfaces* 17, 2 (1987), 78–90.
- [46] MUTH, J. Rational Expectations and the Theory of Price Movements. *Econometrica* 29, 6 (1961), 315–335.
- [47] NICHOLSON, S. Price Ratios in Relation to Investment Results. *Financial Analysts Journal* 24, 1 (1968), 105–109.
- [48] ROSENBERG, B., REID, K., AND LANSTEIN, R. Persuasive evidence of market inefficiency. *Journal of Portfolio Management* 11 (1985), 9–17.

- [49] ROSS, M. Capital Budgeting Practices of Twelve Large Manufacturers. *Financial Management (winter 1986) vol. 15*, 4 (1986), 15–22.
- [50] SALMON, F. Recipe for Disaster : The Formula That Killed Wall Street. *Wired Magazine 17*, 03 (2009).
- [51] SAMORODNITSKY, G., AND TAQQU, M. *Stable non-Gaussian random processes : Stochastic models with infinite variance*. Chapman & Hall.
- [52] SHARPE, W. A Simplified Model for Portfolio Analysis. *Management Science 9*, 2 (1963), 277–293.
- [53] SHREVE, S. *Stochastic Calculus for Finance II : Continuous-Time Models*. Springer Finance. Springer Science and Business Media, Inc., New York, 2004.
- [54] SKLAR, A. Fonctions de répartition à n dimensions et leurs marges. *Publications de l’Institut de Statistique de L’Université de Paris 8* (1959), 229–231.
- [55] SORNETTE, D. *Why Stock Markets Crash : Critical Events in Complex Financial Systems*. Princeton University Press, November 2002.
- [56] SORNETTE, D. Critical market crashes. *Physics Reports 378* (2003), 1–98, arXiv :cond-mat/0301543.
- [57] SORNETTE, D. The Financial Bubble Experiment, 2010. Récupéré le 2010-09-20 sur le site Web de l’ETH Life à l’adresse
[http ://www.ethlife.ethz.ch/archive_articles/100503_prognosenexperiment_nsn/index_EN](http://www.ethlife.ethz.ch/archive_articles/100503_prognosenexperiment_nsn/index_EN).
- [58] SORNETTE, D., AND JOHANSEN, A. Large financial crashes. *Physica A 245*, 3–4 (1997), 411–422, arXiv :cond-mat/9704127.
- [59] SORNETTE, D., AND JOHANSEN, A. Significance of log-periodic precursors to financial crashes. *Journal of Quantitative Finance I* (2001), 452–471, arXiv :cond-mat/0106520.
- [60] SORNETTE, D., JOHANSEN, A., AND BOUCHAUD, J. Stock Market Crashes, Precursors and Replicas. *Journal de Physique I 6* (1996), 167–175, arXiv :cond-mat/9510036.

- [61] SORNETTE, D., WOODARD, R., FEDOROVSKY, M., REIMANN, S., WOODARD, H., AND ZHOU, W.-X. The Financial Bubble Experiment : Advanced Diagnostics and Forecasts of Bubble Terminations Volume II — Master Document, 2010, arXiv :q-fin.ST/1005.5675v1.
- [62] THIELE, T. Om Anvendelse af mindste Kvadraters Methode i nogle Tilfælde, hvor en Komplikation af visse Slags uensartede tilfældige Fejlkilder giver Fejlene en 'systematisk' Karakter. *Det Kongelige Danske Videnskabernes Selskabs Skrifter - Naturvidenskabelig og Matematisk Afdeling* (1880), 381–408.
- [63] WIENER, N. Differential space. *Journal of Mathematical Physics* 2 (1923), 131–174.
- [64] WU, F. The Potts model. *Reviews of Modern Physics* 54, 1 (1982), 235–268.
- [65] ZHOU, W., AND SORNETTE, D. Statistical Significance of Periodicity and Log-Periodicity with Heavy-Tailed Correlated Noise, 2001, arXiv :cond-mat/0110445.
- [66] ZHOU, W., AND SORNETTE, D. Renormalization Group Analysis of the 2000-2002 anti-bubble in the US S&P 500 index : Explanation of the hierarchy of 5 crashes and Prediction. *Physica A* 330 (2003), 584–604.